

DE GRUYTER

REFERENCE

Joan A. Argenter, Jens Lüdtke (Eds.)

MANUAL OF CATALAN LINGUISTICS

MANUALS OF
ROMANCE LINGUISTICS

EBSCO Publishing : eBook Collection (EBSCOhost) - printed on
2/9/2023 9:03 PM via
AN: 135783 ; Joan A. Argenter, Jens Ldtke.; Manual of Catalan
Linguistics
Account: ns335141

Manual of Catalan Linguistics

MRL 25

Manuals of Romance Linguistics

Manuels de linguistique romane

Manuali di linguistica romanza

Manuales de lingüística románica

Edited by

Günter Holtus and Fernando Sánchez-Miret

Volume 25

Manual of Catalan Linguistics

Edited by
Joan A. Argenter and Jens Lüdtke

DE GRUYTER

ISBN 978-3-11-044825-2
e-ISBN (PDF) 978-3-11-045040-8
e-ISBN (EPUB) 978-3-11-044831-3

Library of Congress Control Number: 2019946471

Bibliographic information published by the Deutsche Nationalbibliothek

The Deutsche Nationalbibliothek lists this publication in the Deutsche Nationalbibliografie;
detailed bibliographic data are available on the Internet at: <http://dnb.dnb.de>.

© 2020 Walter de Gruyter GmbH, Berlin/Boston

Cover image: © Marco2811/fotolia

Typesetting: jürgen ullrich typesatz, Nördlingen

Printing and binding: CPI books GmbH, Leck

www.degruyter.com

Manuals of Romance Linguistics

The new international handbook series *Manuals of Romance Linguistics* (*MRL*) will offer an extensive, systematic and state-of-the-art overview of linguistic research in the entire field of present-day Romance Studies.

MRL aims to update and expand the contents of the two major reference works available to date: *Lexikon der Romanistischen Linguistik* (*LRL*) (1988–2005, vol. 1–8) and *Romanische Sprachgeschichte* (*RSG*) (2003–2008, vol. 1–3). It will also seek to integrate new research trends as well as topics that have not yet been explored systematically.

Given that a complete revision of *LRL* and *RSG* would not be feasible, at least not in a sensible timeframe, the *MRL* editors have opted for a modular approach that is much more flexible:

The series will include approximately 60 volumes (each comprised of approx. 400–600 pages and 15–30 chapters). Each volume will focus on the most central aspects of its topic in a clear and structured manner. As a series, the volumes will cover the entire field of present-day Romance Linguistics, but they can also be used individually. Given that the work on individual *MRL* volumes will be nowhere near as time-consuming as that on a major reference work in the style of *LRL*, it will be much easier to take into account even the most recent trends and developments in linguistic research.

MRL's languages of publication are French, Spanish, Italian, English and, in exceptional cases, Portuguese. Each volume will consistently be written in only one of these languages. In each case, the choice of language will depend on the specific topic. English will be used for topics that are of more general relevance beyond the field of Romance Studies (for example *Manual of Language Acquisition* or *Manual of Romance Languages in the Media*).

The focus of each volume will be either (1) on one specific language or (2) on one specific research field. Concerning volumes of the first type, each of the Romance languages – including Romance-based creoles – will be discussed in a separate volume. A particularly strong focus will be placed on the smaller languages (*linguae minores*) that other reference works have not treated extensively. *MRL* will comprise volumes on Friulian, Corsican, Galician, among others, as well as a *Manual of Judaeo-Romance Linguistics and Philology*. Volumes of the second type will be devoted to the systematic presentation of all traditional and new fields of Romance Linguistics, with the research methods of Romance Linguistics being discussed in a separate volume. Dynamic new research fields and trends will yet again be of particular interest, because although they have become increasingly important in both research and teaching, older reference works have not dealt with them at all or touched upon them only tangentially. *MRL* will feature volumes dedicated to research fields such as Grammatical Interfaces, Youth Language Research, Urban Varieties, Computational Linguistics, Neurolinguistics, Sign Languages or Forensic Linguistics. Each volume

<https://doi.org/10.1515/9783110450408-202>

will offer a structured and informative, easy-to-read overview of the history of research as well as of recent research trends.

We are delighted that internationally-renowned colleagues from a variety of Romance-speaking countries and beyond have agreed to collaborate on this series and take on the editorship of individual *MRL* volumes. Thanks to the expertise of the volume editors responsible for the concept and structure of their volumes, as well as for the selection of suitable authors, *MRL* will not only summarize the current state of knowledge in Romance Linguistics, but will also present much new information and recent research results.

As a whole, the *MRL* series will present a panorama of the discipline that is both extensive and up-to-date, providing interesting and relevant information and useful orientation for every reader, with detailed coverage of specific topics as well as general overviews of present-day Romance Linguistics. We believe that the series will offer a fresh, innovative approach, suited to adequately map the constant advancement of our discipline.

Günter Holtus (Lohra/Göttingen)

Fernando Sánchez-Miret (Salamanca)

July 2019

Acknowledgements

The genesis of the present work dates back to 2013, when the series editors proposed to Jens Lüdtke the creation of a *Manual of Catalan Linguistics*. In 1984 Jens had published an introduction to Catalan linguistics for a German-speaking readership (*Katalanisch. Eine einführende Beschreibung*), which provided an overview of the field to meet the growing interest of scholars, students and the general public in the Catalan language. No similar publication existed at the time and the book soon achieved a significant circulation among both students and scholars. By 2013, it had been long out of print, and a new edition would have required a complete revision. This led Jens to accept the series editors' offer, with the aim of providing deeper and updated insights into Catalan linguistics to an even broader readership.

Two questions remained open: whether Jens would be the sole editor of the manual and what would be the language of publication. Jens invited me to get involved early in the project, and I first did so in an advisory role, providing suggestions for the design of the work, topics and authors. As our cooperation on this project grew closer, I received a formal invitation from the publisher and ultimately accepted the role of co-editor. Regarding the language of publication, Jens opted for English, in agreement with the series editors and the publisher, and in keeping with the aim of reaching the largest possible audience.

Unfortunately, Jens Lüdtke passed away on 4th January 2019. After two years of working side by side, I was left to complete our shared mission alone. I was determined to complete the work we had started together. Thirty-five years after his *Einführung*, this *Manual of Catalan Linguistics* pays tribute to Jens' fruitful career.

I would like to thank Jens for his continuous advice and dedication and for never losing hope and good humour even in the most difficult situations. He dedicated all his energy to this project but, tragically, passed away before he could see the manual in print.

My thanks go to the publishing house and the editors of the Manuals of Romance Linguistics series. Without their encouragement and support, I would not have been able to shoulder this new responsibility alone. I would particularly like to thank Dr. Ulrike Krauß for her trust in me and Gabrielle Cornefert for her editorial assistance, as well as Günter Holtus and Fernando Sánchez-Miret, series editors, for their careful supervision of the texts. I must also express our appreciation for the assistance provided by Erin Goldfinch, who revised some of the authors' originals in English and brought them into good stylistic shape. I know that Jens Lüdtke would have expressed similar acknowledgements. I would like to express deep gratitude to Monika Lüdtke, for sharing Jens' material without hesitation and reviewing a preliminary version of the Introduction.

The editors would like to thank the authors of this handbook for their contributions as well as for their cooperation. I am extremely grateful for their encourage-

<https://doi.org/10.1515/9783110450408-203>

ment – in the most critical moment of creating this work – to accomplish the common goal, that is, to bring this *Manual of Catalan Linguistics* project to fruition.

Joan A. Argenter
Barcelona, July 2019

Table of Contents

Joan A. Argenter

0 Introduction — 1

Joan F. Mira

1 Languages, Cultures, Nations: A History of Europe — 17

Joan Julià-Muné

2 History of Catalan Linguistics — 37

Language Description

Xavier Lamuela

3 Spelling — 81

Nicolau Dols

4 Phonology, Phonetics, Intonation — 101

5 Morphosyntax — 129

Mar Massanell i Messalles

5.1 Word Classes, Inflectional Categories and Paradigms — 129

Gemma Rigau and Manuel Pérez Saldanya

5.2 The Simple Sentence — 165

Manuel Pérez Saldanya and Gemma Rigau

5.3 The Complex Sentence — 211

Ingo Feldhausen and Xavier Villalba

5.4 Modality and Information Structure: Focus, Dislocation, Interrogative and Exclamatory Sentences — 247

Jaume Mateu

5.5 Lexicalized Syntax: Phraseology — 271

Maria Josep Cuenca

6 Pragmatics and Text Linguistics — 287

7 Lexicon — 311

Josep Martines

7.1 General Lexicon — 311

Jens Lüdtke

7.2 Word-Formation — 351

8 Variation and Varieties — 371

Mar Massanell i Messalles

8.1 Dialects — 371

Miquel Àngel Pradilla Cardona

8.2 Social and Functional Variation in Catalan — 397

Joan Soler Bou

9 Language Corpora — 421

Language History

Philip D. Rasico

10 Early Medieval Catalan — 437

Antoni Ferrando

11 The Growth and Expansion of Catalan (1213–1516) — 471

Miquel Nicolás

12 The Origins of Modern Catalan: Cultural and Linguistic Evolution — 485

Jenny Brumme

13 *Renaixença* — 497

14 Towards Language Institutionalization — 517

August Rafanell

14.1 The Language Reform, the Institut d'Estudis Catalans and the Work of Pompeu Fabra — 517

August Rafanell

14.2 From Pompeu Fabra to the Present Day: Language Change, Hindrance to Corpus and Status Planning — 545

Albert Turull

15 Onomastics: Personal Names and Place Names — 561

Montserrat Bacardí and Joaquim Mallafrè

16 Translation — 581

Catalan Today

Joan A. Argenter

17 Languages in Contact: A Sociocultural Approach — 597

F. Xavier Vila

18 Language Demography — 629

Eva Pons

19 Language Law and Language Policies — 649

F. Xavier Vila

20 Teaching and Learning of Catalan — 669

Oriol Camps Giralt and Aina Labèrnia Romagosa

21 Catalan in the Mass Media: The Rise of Stylebooks — 683

M. Teresa Cabré and M. Amor Montané

22 Terminology and Neology — 693

Emili Boix-Fuster and Kathryn A. Woolard

23 Language Ideologies in Society — 709

Joan Pujolar

24 Migration in Catalonia: Language and Diversity in the Global Era — 723

August Bover i Font

25 Catalan Worldwide — 739

XII — Table of Contents

List of Contributors — 751

Index — 755

Joan A. Argenter

0 Introduction

1 Preamble

This introductory text to the *Manual of Catalan Linguistics* was intended to be a collaborative effort between the two co-editors, much like the majority of the planning and execution of the editorial work for this handbook has been. Regrettably, the intended collaboration cannot now be achieved: Jens Lüdtke passed away before the task was undertaken.

Even so, the authorship of this introduction is not easily defined. Both Jens and I drafted parts of it at an earlier stage in the editing process: we tried to find common ground on any controversial matters. We had planned to write the final text together, but it was left to me to do so alone. Nevertheless, Jens' voice can still be heard in these pages.¹

2 Basic features

This editorial undertaking started with the observation that no comprehensive overview of Catalan linguistics was available, neither in Catalan nor in any other language. To fill this gap and cater to both beginner and advanced students of Catalan, both in Catalan-speaking regions and elsewhere, the *Manual of Catalan Linguistics* combines an in-depth presentation of elementary principles with the outline of more complex questions and new research fields. Thus, the texts have been prepared with scientific rigour and include, as much as possible, current research in the fields in question. The language is intended to be clear and accessible to the general reader. The manual is intended to support students of Catalan in their studies and introduce specialists of other languages to this field, in particular scholars of the Romance languages.

The general editors of the series *Manuals of Romance Linguistics* have stressed their interest and concern for “minority languages” and the reader's assumption may be that Catalan belongs to this category. It is not my intention here to argue whether this claim is self-evident. Many scholars agree that Catalan is a singular case, given its relatively robust demography and status in terms of social use and prestige, political recognition and development of the language. At the same time, Catalan is still a subordinated language. Minorities are usually defined by scholars, politicians or

¹ In writing this introduction I have used some of Jens' ideas or even his own wording as they appear in earlier fragmentary drafts. The responsibility for this text, however, lies solely with the author.

laymen belonging to a majority who identify themselves with a “majority language” or community. The label “small” languages and communities, as opposed to “major” ones, remains in use in recent literature from a variety of theoretical perspectives (Dorian 2014; Pietikäinen et al. 2016). Currently, Catalan sociolinguists define Catalan as a ‘medium-sized European language’ (Boix-Fuster/Farràs 2013; Milián-Massana 2012; Vila/Bretxa 2013), which descriptively and prospectively defines a state of affairs which entails certain consequences.²

3 Plan and structure of the manual

The editors agreed to organise this *Manual of Catalan Linguistics* thematically rather than by discipline or theory. In a way, this gave us the opportunity to gather together authors of different disciplines and to favour a variety of views on related topics as well as to make it easier for them to tackle the task assigned to them (obviously, the line between theme and discipline is rather thin in some cases, as these terms can be treated synonymously). Chapters on the structure of the language are organised according to the level of analysis. There is no single chapter on sociolinguistics; this may appear to some an incomprehensible omission given the Catalan sociolinguistic setting and the exemplary autochthonous production in the field, but the subject is so interdisciplinary in nature that it is distributed across several chapters. Indeed, the themes dealt with in the book have been assigned to a representative group of highly renowned scholars, including linguists, philologists, sociolinguists, anthropologists and many others, whose expertise in their chosen topic or field is beyond question. In principle, the thematic approach could lead to authors beginning their texts *in medias res* rather than first of all laying out some preliminary theoretical foundations – although they were given free rein to follow either approach. On the other hand, in a few cases an author’s choices regarding his or her chapter has motivated the editors to move his or her chapter forward or back in the formerly planned structure of the work, either from the descriptive to the historical block or vice versa, according to where the text was best suited after reviewing the manual as a whole.

For instance, terminology – which we shall take as an example for the sake of argument – is part and parcel of the lexicon of a language, together with general lexicon (7.1 General Lexicon). However, the development of Catalan terminology in the last three decades and a half has been led by Termcat, a centre created in 1985 as a consortium between the Generalitat de Catalunya (Catalan Government) and the Institut d’Estudis Catalans (IEC), the institution responsible for corpus language plan-

² Medium-sized language communities range from “a lower limit of one million speakers and an upper limit of 25 million” – a figure based on theoretical and strategic criteria: “In practical terms, for instance, this limit for MSLCs includes all the official languages of the European Union that are traditionally considered to be small” (Vila/Bretxa 2013, 7).

ning. The aim of this strategic centre is to coordinate terminological activity in Catalan, the creation of terminological products and the standardisation of neologisms, to ensure the availability of Catalan terminology in all sectors of activity and knowledge and to encourage its use. It coordinates with and gives advice to other terminology-creating bodies and provides specialists and the general public with terminology resources in various formats. Termcat works together with the IEC and professionals in a number of fields. Not least, “Terminology” as an applied discipline has developed recently as a response to denominative and communicative needs. In summary, terminology and neology are very closely linked with modern life and with language planning, and therefore with the current situation of the language. Therefore, it should not come as a surprise that the authors’ treatment of the subject in the chapter on this topic revolves around these issues (722 Terminology and Neology), i.e. the creation of infrastructure for the management of Catalan terminology and the development of the theoretical underpinnings for a new discipline. The authors’ choices and their considerations motivated us to shift their chapter from the descriptive section of the manual to the section on the current situation. All this brings us to the structure of the volume.

Following this introduction to the *Manual of Catalan Linguistics*, the relationship between languages, cultures and nations in Europe – framing the ideology of language and national identity – is examined (71 Languages, Cultures, Nations: A History of Europe), as well as the history of Catalan linguistics (72 History of Catalan Linguistics). The Herderian ideology of language and nation – and its practical adoption and implementation by modern nation-states – shaped the environment in which European “minority languages” evolved while at the same time providing minorities with a model to emulate. Humboldt’s notion of a language as a specific historic and cultural phenomenon, his view of the relationship between a language and the people who speak it and his conviction of the link between the inner form of a language and its speakers’ way of thinking were highly influential among nationalist movements in Europe throughout the 19th and 20th centuries. This influence also permeated the politics of Catalanism and its cultural and intellectual manifestations (Prat de la Riba 1906; 1908).³ The discourse of language and national identity has been passed down from generation to generation up to the present day. Nowadays,

3 Prat de la Riba (1870–1917) was a regionalist right-wing politician who imbued the social movement of Catalanism with a political programme and a nationalist ideology. He presided over the *Mancomunitat de Catalunya* ‘Commonwealth of Catalonia’ (1914–1917) – a trial of a unified highly limited self-government of Catalonia, overcoming its division in four administrative provinces –, and was the founder of the *Institut d’Estudis Catalans* (1907). Even a left-wing federalist such as Valentí Almirall (1841–1904), who considered not language, but a ‘community of moral and material interests’ to be the foundation of a people’s personality, endorsed the Humboldtian idea of the link between a language and the character of a people, between expression and conception, between language and thinking and feeling (Almirall 1886, 89). Almirall’s represents the first systematic exposition of Catalanism as a political movement twenty years before Prat’s.

however, while this ideology maintains some of its social relevance, it has been subjected to a certain amount of critical analysis, while other ideologies have emerged from a new, complex society in a global world (↗23 Language Ideologies in Society; ↗24 Migration in Catalonia: Language and Diversity in the Global Era).

Beyond that, this handbook is divided into three parts: (1) a description of Catalan, according to the viewpoint of contemporary synchrony on the levels of phonology and phonetics, morphology and syntax, the lexicon, language variation and varieties – without forgetting the pragmatic and textual dimensions of the language –; (2) the evolution of the language, or the diachronic viewpoint, as well as the social history of the language; (3) the present situation – or “Catalan today” – , which unites both perspectives and highlights the role and usage of the language in society, including the social relevance of language practices, language ideologies, language (particularly educational) policies, language law and the politics of language within a specific linguistic demography. The impact of recent immigration from third countries is dealt with, from both demographic and ethnographic perspectives.

By the very nature of its subject, certain chapters cover a systematic and quasi-complete description of its matter (↗4 Phonology, Phonetics, Intonation), while others are constrained to select a few topics to deal with – and show, for instance, how language structure and use interact with each other in interesting and language-specific ways (↗6 Pragmatics and Text Linguistics).

An area such as grammar – traditionally including morphology and syntax – demanded a rather homogeneous approach, something that has been achieved to a large extent. This goal, however, does not preclude that authors from different stance provide an overall nuanced presentation of distinct subareas of this field. Thus, formal approaches (↗5.1 Word Classes, Inflectional Categories and Paradigms; ↗5.2 The Simple Sentence; ↗5.3 The Complex Sentence; ↗5.4 Modality and Information Structure: Focus, Dislocation, Interrogative and Exclamatory Sentences) go hand in hand with semantic-based approaches such as from-content-to-form analysis (↗7.2 Word-Formation) or cognitive grammar (↗5.5 Lexicalized Syntax: Phraseology). All in all, a rather descriptive picture of the field emerges, with technicalities sufficiently explained or reduced to a minimum.

As a matter of course, the study of the lexicon of Catalan belongs to its description – although words have their own history and social development increases the lexical repertoire to satisfy new expressive needs. Both the social and the cognitive mechanisms to expand vocabulary – among which metaphor and metonymy play a role – should be considered together (↗7.1 General Lexicon). Today research in this and other language areas take advantage of the constitution of digitized resources, either yielding old lexicographic works in electronic form or developing repositories of texts permitting distinct operations of information retrieval (↗9 Language Corpora).

Part (2), devoted to the evolution and history of Catalan, encompasses the emergence and formation of Catalan, from its Vulgar Latin origins, and its transformation into an individuated Romance language (↗10 Early Medieval Catalan) as well as its

social and cultural history in the flourishing medieval period (↗11 The Growth and Expansion of Catalan (1213–1516)), its later decay in the modern period (↗12 The Origins of Modern Catalan: Cultural and Linguistic Evolution), its revival in the nineteenth century (↗13 *Renaixença*) and the Language Reform in the early twentieth century (↗14.1 The Language Reform, the Institut d'Estudis Catalans and the Work of Pompeu Fabra) up to the attempt to ban it from public life or definitively eradicate it by two dictatorships in the twentieth century (↗14.2 From Pompeu Fabra to the Present Day: Language Change, Hindrance to Corpus and Status Planning) – and well before (↗20 Teaching and Learning of Catalan).

The second and third parts are more closely linked together than is apparent from the table of contents. For instance, the authors in the third section address the historical aspects of their subjects which do not appear, or do so only marginally, in the chapters on the history of the language. Other chapters in this block concern crucial aspects in the spread and consequential development of Catalan as an official and public language (↗20 Teaching and Learning of Catalan; ↗21 Catalan in the Mass Media: The Rise of Stylebooks). A view on Catalan outside its homeland or Catalan abroad closes the volume (↗25 Catalan Worldwide).

4 Webs of words: translation and literary language. A case in point

This section of the introduction is devoted to some specificities of the study of Catalan linguistics that we would like to draw the reader's attention to. These have to do with the interaction between the standardisation of Catalan and the development of 'literary language' – or to put it in technical terms, the interplay between language codification and language elaboration – and with the way in which linguistic disciplines can provide insight into this development.

Ideally, multiple different chapters would have addressed issues such as the impact of translation on 'literary language' (the term used by Pompeu Fabra and many contemporaries, including Ferdinand de Saussure and members of the Prague Linguistic Circle) – or vice versa – or its impact on 'standard language' (a non-equivalent prevalent term today) – or vice versa.⁴ Moreover, it would have been useful

4 Pompeu Fabra (1868–1948), a chemical engineer turned grammarian and lexicographer, was the man who spearheaded the standardisation of modern Catalan from the IEC in the first third of the 20th century (↗14.1 The Language Reform, the Institut d'Estudis Catalans and the Work of Pompeu Fabra). Aside from his technical training and a brief period of technical teaching in Bilbao, Basque Country, he devoted his life to the study of Catalan and wrote several grammars, a dictionary, teaching material, essays and notes on language issues in addition to a large amount of civic literature (Fabra 2005–2012). He gained the respect and trust of Catalan politicians, intellectuals, writers and society at large. Driven by his civic-mindedness, he was a key agent in Catalan intellectual, academic and political life. In the

to see examples of outcomes in term formation – specifying the criteria applied in several different cases – or examples of different choices of standard language from mass media, and so on, which would have allowed us to observe language trends. However, this goal could not be easily and coherently met in chapters that had to prioritise their subject matter according to the authors' choices. The chapter on translation takes a historical approach, providing an excellent review of Catalan translations from the Middle Ages up to the present day (716 Translation). To illustrate my point, let us compare the translation of a biblical verse from the *Song of Songs* in three different Catalan versions (covering more or less the span of a century).⁵

<i>Al nadó de la cerva, a la gazela, lo meu Aimat se sembla.</i> (Jacint Verdaguer 1907) ⁶	<i>El meu amat és igual que la gasela O que el cabrit de les cérvoles.</i> (Carles Riba 1918) ⁷	<i>A una gasela, a un cervatell s'assembla l'estimat.</i> (Narcís Comadira/Joan Ferrer 2013) ⁸

"My beloved is like a gazelle or a young stag."

Leaving aside spelling issues – Verdaguer, whose work greatly contributed to the development of modern Catalan, was writing before the Language Reform (714.1 The Language Reform, the Institut d'Estudis Catalans and the Work of Pompeu Fabra); Riba's translation came after the Spelling Reform (IEC 1913; IEC 1917; 73 Spelling) and issued the same year that the first official grammar appeared (Fabra 1918); the

aftermath of the Spanish Civil War, he went into exile in France and served as a 'Counsellor' of the Catalan Government in exile (1945–1948). Eventually, he settled in French Catalonia (*Catalunya del Nord*) and died on Catalan soil as was his will.

5 I have chosen translations of the *Song* by Catalan poets. A comparative study of different versions of the Bible translations by churchmen, monks or Hebraists would bring us to similar conclusions.

6 Posthumously published (Verdaguer 1907). Jacint Verdaguer (1885–1902) or Mossèn Cinto Verdaguer – the title betrays his position as a priest – was a very well-known figure and is usually referred to by his title and nickname in an expression of both power and solidarity (Brown/Gilman 1960). Verdaguer was a great late Romantic poet within the *Renaixença* movement (713 *Renaixença*). He cultivated folk lyrical poetry, epic poems and prose.

7 Carles Riba (1893–1959) was one of the greatest Catalan poets, writers and translators of the 20th century. He was a specialist in Classics. As a Hellenist, his versified translation of Homer's *Odyssey* is a landmark of Catalan literature. He also translated tragedies of Sophocles and Euripides. He translated other classical authors from both Greek and Latin as well as modern writers from German (Grimm, Hölderlin, Rilke), English (Poe), French (Bédier), Italian and Modern Greek (Kavafis). He presided over the Philological Section of the IEC and undertook the task of editing Fabra's *Diccionari general de la llengua catalana* (Fabra 1954), which was subject to censorship under the Francoist regime (for a reproduction of censored fragments, see Institut d'Estudis Catalans. Secció Filològica 1990).

8 Narcís Comadira (1942–) is a Catalan poet, translator and painter. Joan Ferrer (1960–) is a Catalan philologist specialised in Semitic languages, mainly Hebrew and Aramaic. He collaborated closely with Joan Coromines (1905–1997), especially in the edition of *DECat* vol. 10.

Comadira/Ferrer translation (the outcome of the collaboration of a poet and a Hebraist) is quite recent –, these verses differ in several relevant lexical and grammatical aspects. One key term in the original Hebrew is conveyed by a different form in each version: *Aimat*, *amat*, *estimat* ‘beloved’. *Aimat* is an Occitan form – occasionally used as a poetic word in Catalan –; *amat* is the corresponding Catalan form (past participle of *amar* ‘to love’ and also a noun);⁹ *estimat* is the most commonly used term in Catalan, from the verb *estimar*, the common everyday word meaning ‘to love’, with no specific poetic connotations.¹⁰ We therefore observe an evolution towards unmarked, universally used terms. Another Hebrew term resulting in three distinct translations is ‘son of doe’: *nadó de la cerva*, *cabrit de les cérvols*, *cervatell*, lit. ‘doe’s newborn’, ‘doe’s kid’, ‘small deer’. Verdaguer’s *cerva* appears as the normative *cérvols* in Riba – who confuses or extends the meaning of *cabrit* ‘kid’ – while the most recent translation renders the meaning synthetically with a diminutive form, *cervatell*.¹¹ These examples reflect what was at stake at the time each of these translations was

9 The word reminds us of Ramon Llull’s (1232–1315) poetry in his *Llibre d’Amic e Amat* ‘Book of the Lover and the Beloved’. What was in common use in Llull’s time today sounds like an old-fashioned poetic word.

10 The traditional medieval form *aesmar*, meaning ‘to estimate’ (13th century; e.g. Llull, the *Great Chronicles*) declined in use in favour of the Latinate form *estimar* with the same meaning. Later on, *estimar* progressively took on the meaning of the traditional verb *amar* ‘to love’ up to the point that the former substituted the latter entirely. When exactly this semantic change took place has been the subject of much debate. In all likelihood, *estimar* ‘to love’, is not as recent as some have assumed and was spread at the beginning of the 17th century, having previously passed through a series of semantic changes (‘to value’, *estimar més* ‘to prefer’) (*DECat* III, 595–599, s.v. *esma*; 77.1 General Lexicon).

11 Verdaguer also used the forms *cérvol* and *cérvola* elsewhere – the only pre-normative occurrences according to the *CTILC*. All the other occurrences in this corpus are from 1914 onwards, following the Spelling Reform. *Cerva* is the feminine of *cervo*. This word perhaps was perceived as a loanword from Spanish – because of its (unaccented) final -o – by Fabra, who did not include the form in his dictionary (Fabra 1932) – about to become the prescriptive dictionary –, despite it being a historical Catalan form since the early days of the language. Instead, Fabra included the form *cérvol*, which Riba obediently used in his version. Actually, the word displays a certain degree of polymorphism across the entire area of Catalan. An actual case of Spanish interference was *ciervo* – not because of its final vowel, but its diphthong. Many words ending in -o were suspect, since this is a common feature of Spanish, but some indisputably Catalan words, like *ferro* ‘iron’, *llobarro* ‘sea bass’, *musclo* ‘mussel’, and all the first person indicative verb forms of Central – and standard – Catalan finish with -o (pronounced [u] because of the vowel reduction rule): *çaço*, *temo*, *dormo* (from *caçar* ‘to hunt’, *témer* ‘to fear’, *dormir* ‘to sleep’). The classical – and still the Balearic – forms were, however, *caç*, *tem*, *dorm*. The non-etymological final -l, as in *cérvol*, has been used to create apparent learned words, mostly hypercorrections, or to give a Catalan form to loanwords. Coromines (*DECat* II, 687–689, s.v. *cervo*) attests to the fact that in a list of amendments and additions to the dictionary drafted by Fabra himself and sent, from his exile in French Catalonia, to the Philological Section of the IEC in 1948, ‘Fabra entrusted <their members> to substitute the dictionary entries *cérvol* (o *cervo*) and *cérvola* for *cervo* (o *cérvol*) and *cerva* (o *cérvola*)’. Be that as it may, these alleged changes did not go ahead and the latter entries did not appear in the second edition of the dictionary (Fabra 1954). Currently, the official dictionary (*DIEC*) contains both forms as independent, though related, entries, with *cervo* cross-referring to *cérvol*.

produced: the shift from *cerva* to *cérvoles* is a matter of prescriptive norms, while the shift from *amat* to *estimat* results from a desire to bring literary language closer to everyday language. Verdaguer still uses the medieval (and current Northwestern Catalan) article *lo* in his version, while in the other texts the current form *el* (or *l'* before vowel) is used. Distinct forms in the three translations are also *se sembla*, *és igual*, *s'assembla* '(he) looks like'. Indeed, while the first and last coincide in form and meaning – though not in spelling –, the second means 'is equal to/equals'. The obvious difference in spelling between Verdaguer's and Comadira/Ferrer's versions plausibly implies different analyses of the verbal sequence or of the lexical form of the verb in moving from orality to literacy (*semblar* 'to seem' is an actual form of the language as is *assemblar-se* 'to look like': in Verdaguer's mind two forms fused into one). Ferrer/Comadira's is the only normative form since the Reform. The order of the lines with regard to the original Hebrew has been reversed in Riba's version.¹² Thus, on the basis of a few translated lines produced at different times, we can observe the emergence of the complex relationship between lexical choices, verbal repertoire, the standardisation process and literary language development.

5 The Catalan-speaking community

At this point it is necessary to highlight a problem that distinguishes the Catalan language from others, at least in the European context. Since the Spanish constitution forbids the unification of the provinces and regions which have Catalan as their common language, any common enterprise – and this handbook is no exception – is an act of self-assertion. According to the Spanish Constitution, Castilian – the only language it mentions – is the official state language, while 'the other Spanish languages' – a clumsy phrase indeed –, of which Catalan is presumably one, are official in their respective Autonomous Communities (the term for the Spanish administrative regions). Today, the power to determine policy with regard to these 'other Spanish languages' is transferred to the respective regions. As a consequence, it may well be that no unitary language policy is applied to speakers of one and the same language when this language crosses borders, as is the case with Catalan. As a matter of fact, there are inequalities in the recognition, implementation and enforcement of speakers' linguistic rights depending on which side of the border they reside in. Also, decisions may be made that harm the unity of the language and its standard form – even though the state, for its part, is obliged to protect this language (19 Language Law and Language Policies).

¹² Verdaguer translated these verses from the Latin *Vulgata*. Riba and Comadira/Ferrer translated them from the original Hebrew (Ferrer/Feliu 2012).

Nonetheless, Catalan linguists have a global view of their linguistic field, unconstrained by political or administrative borders, and present the facts accordingly (↗8.1 Dialects; ↗8.2 Social and Functional Variation in Catalan – chapters on grammar or language history could be added to make a point –). By contrast, although sociolinguists share this global view, sometimes the choice of subject matter requires that these borders be taken into account in an otherwise global presentation of empirical facts, because they reveal an uneven state and evolution across borders (↗8.2 Social and Functional Variation in Catalan; ↗17 Languages in Contact: A Sociocultural Approach; ↗18 Language Demography; ↗19 Language Law and Language Policies). Likewise, the awareness of the unity of Catalan and of its entire domain underlies the comprehensive practices aimed at the determination of a compositional and polymorphic model of standard Catalan, thus integrating the most relevant features of the main Catalan dialects in one prescriptive corpus (↗8.2 Social and Functional Variation in Catalan; IEC 1990; IEC 1992; *DIEC* 2007; *GIEC* 2016).¹³

Yet, acts of self-assertion do not arise from institutional politics or legal constraints alone. Everyday language use and interaction constitutes a power domain and, indeed, control over a language minority is not only exerted through legal and political oppression, but also through everyday interaction – a space where the struggle for dominance and resistance emerges through language use patterns mediated by language ideologies. Hence, everyday interaction and practices also merit scrutiny (↗17 Languages in Contact: A Sociocultural Approach).

As has been asserted above, the scope of this handbook is the entire domain of Catalan stretching from Andorra and the northernmost Catalan-speaking trans-Pyrenean region in France to the southernmost part of the Valencian Country; from the Catalan-speaking Strip in Aragon to the Balearic Islands and beyond, to the town of l'Alguer, a Catalan-speaking enclave in Sardinia. Occasionally, groups other than “strictly territorial” ones have been taken into consideration, such as Catalan gypsies in France – ethnic groups who identify themselves as Catalan gypsies and claim Catalan as their ethnic language (↗17 Languages in Contact: A Sociocultural Approach; ↗8.1 Dialects). While political and administrative borders are relevant, mainly as regards sociolinguistic setting and status planning, political or administrative and linguistic borders are not one and the same. The extent to which the linguistic

13 The term ‘compositional’ was first defined in Catalan sociolinguistics by Polanco (1984), within a broader precise schema of codification types. By this term, he meant a very precise type of codification, including not only elements from different varieties but also a single codification centre and the creation of a single norm, although this norm may allow regional applications. Haugen (1968) also used the same term (referring to Nynorsk), but in a much looser sense, to refer to a composite norm derived from a set of related (Norwegian) dialects. Polanco’s theoretical model proved to be quite useful in explaining the main features of the Catalan standardisation process (and indeed other types of norm creation), including not only Fabra’s codification, but also its subsequent adaptations and developments.

identity across this area corresponds with ethnic or cultural identities is a matter for research – and occasionally a contentious one in terms of individual attitudes and beliefs. Thus, the entire Catalan-speaking community, including its “margins”, is considered, whatever its internal cohesion or foreign recognition.

That said, it should be noted that Catalan linguistics and sociolinguistics are not detached from the crisis of the notions of “language” and “native speaker”. The crisis over the notion of a language as a bounded system and its substitution by the notion of a linguistic repertoire (since the early days of Gumperz 1964; Blom/Gumperz 1986, 1972) was assumed by Catalan sociolinguists as well as the crisis over a native speaker-based linguistics or the notion of “native speaker” itself. It can even be asserted that the crisis over the notion of the native speaker and the focus on the “new speaker” paradigm has emerged – albeit not exclusively – from the very heart of Catalan sociolinguistics. Research projects have been conducted on “new speakers” (O’Rourke/Pujolar 2013; O’Rourke/Pujolar/Ramallo 2015) and research on multilingual repertoires is on the rise.¹⁴ The “new speakers” notion itself, however, appears to be equivocal, since it covers a range of sociolinguistic trajectories, e.g.: (i) migrants: either (ii) migrants from economically deprived regions in Spain in the 20th century or (iii) third-country migrants arriving in Catalonia in the 21st century. For one thing, these two categories are quite distinct because of the origin, native languages and the cultural and religious proximity or distance between migrants and locals. Moreover, (iv) migrants may have learnt Catalan in institutional settings as a result of top-down policies or (v) in natural social settings (and the result may differ in terms of competence and local legitimacy). A different case is that of (vi) speakers living in the Catalan-speaking area who “recovered” their ancestors’ traditional language, which they did not learn directly from their parents or relatives, by way of reversing language shift (RLS, Fishman 1991) or revitalising programmes; other cases could be added to the list. However, taking a critical look at native speaker and native speech communities successfully brought up the sensitive issue of what the “legitimate code” is and to what extent new speakers are recognised as “legitimate speakers” of this code in certain contexts. It also points to the fact that individuals in a society have distinct linguistic trajectories and make their own choices; in a way, they become agents of language planning on the ground. Analysing the watershed moments or key points in life at which Catalan – whether acquired as a result of institutional or natural learning or of reversing language shift – is actually put to use sheds light on how individuals and social groups develop their relationships to components of their repertoire and how they construct their sense of belonging to a chosen speech

14 Joan Pujolar (UOC – Open University of Catalonia –, Barcelona), Fernando Ramallo (University of Vigo, Galicia), Bernadette O’Rourke (Heriot-Watt University, Scotland), Estibaliz Amorrotu (Universidad de Deusto, Basque Country) and their teams have pioneered research in the field. O’Rourke and Pujolar have coordinated a European research network entitled “New Speakers in a multilingual Europe: Opportunities and challenges” (2013–2017).

community (Pujolar/Puigdevall 2015). This process raises issues not only related to the legitimacy of new speakers, but to claims of authenticity, authority and ownership among traditional speakers (↗23 *Language Ideologies in Society*).

6 Editors' notes to the reader

This manual claims to be representative of Catalan Linguistics as cultivated mainly by the Catalans, Valencians and Balearics themselves, and some non-Catalans, within the limits of the number of authors, but it should be clear from the outset that the approaches presented here cannot be homogeneous by virtue of the very representativeness we wish to guarantee. The editors have not tried to smooth out these differences, but have instead chosen to reflect the dynamic nature of this field. Besides, it should not come as a surprise that, aside from “foreign” authors, the majority of the contributors belong to a generation of scholars who helped to establish and shape Catalan as an academic discipline at universities after the Spanish Transition or who were shaped by them.

6.1 On Catalan orthography

While this handbook was being drawn up, the Institut d'Estudis Catalans published a new spelling guide (*Ortografia catalana*, IEC 2017), which contains small changes with regard to the spelling prevailing for more than a century – not all of these changes have been introduced currently, but were approved decades ago while others are based on recent decisions. A period of four years has been set before the new spelling rules come into force – generally, however, government administration, the school system, publishing houses and newspapers are applying it from the outset. Early versions of texts for the chapters in this *Manual* reached the editors before these new spelling rules were issued while others arrived afterwards. The editors suggested that the authors adapt fragments or examples in Catalan in their texts to the new orthographic rules. However, there may be some variation resulting from the authors' choices or the lack of comprehensive adjustments by revisers.

6.2 On quotations and quotation marks

A quick note on quotations and the discriminative use of quotation-marks: as a matter of principle, original quotations from English or German appear within “double quotation marks”. Original Catalan quotations translated into English appear within ‘single quotation marks’ in order to make it clear to the reader which is the original language. Generally, quotations from the major Romance languages, including

French, Spanish, Portuguese and Italian, are also presented in the original language within “double quotation marks”, since generally they are not translations. When more than one language appears in mixed texts, the selective use of letter types is stated in the chapter.

6.3 On place names

Often assigned a peripheral place in mainstream linguistics, onomastics is, nonetheless, a relevant dimension of a language in terms of both collective and personal identities as well as language landscape (≥15 Onomastics: Personal Names and Place Names). For some reason authoritarian regimes or highly centralized nation-states often distort the personal names of the members of subordinated minorities as well as minorities’ place names when these are in a language other than the one of those who hold power. Likewise, one of the earlier symptoms of irreversible language decay in language death settings is the adoption of majority names by speakers of the receding language (Dressler 1987). Since the recovery of self-government institutions in the Catalan-speaking regions in Spain this trend has been reversed there and autochthonous names recovered.

We make a daring choice regarding the graphic presentation of place names, which departs from common use in English. Initially, we recommended using the current English place and proper names, but could not avoid acknowledging that this practice is contrary to the editors’ claim to be presenting the Catalan language from the speakers’ point of view. Whenever the form of an English name is firmly established, we advocate the general use. However, in many cases the Spanish, French and Italian names are neither sufficiently well-known nor do they reflect appropriately the otherness of the Catalan Countries or the way their inhabitants see themselves. This option may have given rise to unusual blending forms in derivatives, such as *Alguerese*, the name of the local Catalan spoken in l’Alguer (It. Alghero), with *-gu-* rather than *-gh-*. We hope that this will not cause insurmountable trouble to the reader. The names and/or their equivalents used in this manual are presented below:

Alacant	Sp. Alicante
Alacantine	Cat. alacantí, Sp. alicantino
l’Alguer	It. Alghero
Alguerese	Cat. alguerès, It. Algherese
Balearic	Cat. balear
Balearic Islands	Cat. Illes Balears, Sp. Islas Baleares
Castelló de la Plana	Sp. Castellón de la Plana
Catalonia	Cat. Catalunya
Central Catalan	Cat. català central
Cotlliure	Fr. Collioure

Eastern Catalan	Cat. català oriental
Eivissa	Sp. Ibiza
Eivissenc	Sp. Ibicenco
Elna	Fr. Elne
Elx	Sp. Elche
Formeteran/Formenterenc	Cat. Formenterenc
Girona	Sp. Gerona
Lleida	Sp. Lérida
lleidatà	Sp. leridano
Maó	Sp. Mahón
Majorca	Cat. Mallorca
Majorcan	Cat. mallorquí
Minorca	Cat. Menorca
Minorcan	Cat. menorquí
Northern Catalan	Cat. català septentrional
Northern Catalonia	Cat. Catalunya del Nord
Northwestern Catalan	Cat. català nord-occidental
Perpinyà/Perpignan	Fr. Perpignan
Pityusic Islands	Cat. Illes Pitiüses, Sp. Islas Pitiusas
Rosselló	Fr. Roussillon
Rossellonese	Cat. rossellonès, Fr. roussillonnais
Sardinia	Cat. Sardenya, It. Sardegna
Sardinian	Cat. sard, It. sardo
Valencia/València	Cat. València/Sp. Valencia
Valencian	Cat. valencià
the Valencian Country	Cat. País Valencià
Western Catalan	Cat. català occidental

7 Bibliography

- Almirall, Valentí (2009, ¹1886), *Lo catalanisme. Motius que'l·llegitimen. Fonaments científics y solucions pràcticas*, Barcelona, Llibreria de Verdaguer i Llibreria de López. Facsimile edition: Barcelona, Centre Excursionista de Catalunya.
- Blom, Jan-Petter/Gumperz, John J. (1986, ¹1972), *Social Meaning and Linguistic Structure. Code-Switching in Norway*, in: John J. Gumperz/Dell H. Hymes, *Directions in Sociolinguistics. The Ethnography of Communication*, Oxford, Blackwell, 407–434.
- Boix-Fuster, Emili/Farràs, Jaume (2013), *Is Catalan a medium-sized language community too?*, in: F. Xavier Vila (ed.), *Survival and Development of Language Communities*, Bristol, Multilingual Matters, 157–178.
- Brown, Roger/Gilman, Albert (1960), *The Pronouns of Power and Solidarity*, in: Thomas A. Sebeok (ed.), *Style in Language*, Cambridge (Mass.), M.I.T. Press, 253–276.
- Comadira, Narcís/Ferrer, Joan (ed. and transl.) (2013), *Càntic dels càntics de Salomó*, Barcelona, Fragmenta.

- CTILC = Institut d'Estudis Catalans, *Corpus Textual Informatitzat de la Llengua Catalana*, <https://ctilc.iec.cat/> (last accessed 19.06.2019).
- DECat = Coromines, Joan (1980–2001), *Diccionari etimològic i complementari de la llengua catalana*, 10 vol., Barcelona, Curial.
- DIEC = Institut d'Estudis Catalans (²2007), *Diccionari de la llengua catalana*, Barcelona, Enciclopèdia Catalana/Edicions 62, <https://dlc.iec.cat/> (last accessed 17.06.2019).
- Dorian, Nancy C. (2014), *Small-Language Fates and Prospects. Lessons of Persistence and Change from Endangered Languages. Collected Essays*, Leiden/Boston, Brill.
- Dressler, Wolfgang U. (1987), *La mort de les llengües*, Límits. Revista d'assaig i d'informació sobre les ciències del llenguatge 2, 87–97.
- Fabra, Pompeu (1918), *Gramàtica catalana*, Barcelona, Institut d'Estudis Catalans.
- Fabra, Pompeu (1932), *Diccionari general de la llengua catalana*, Barcelona, Catalonia.
- Fabra, Pompeu (1954), *Diccionari general de la llengua catalana*, Barcelona, López Llausàs.
- Fabra, Pompeu (2005–2012), *Obres completes*, ed. Jordi Mir/Joan Solà, 9 vol., Barcelona, Proa.
- Ferrer, Joan/Feliu, Francesc (2012), *La traducció del Càntic dels Càntics de Carles Riba*, Tamid: Revista Catalana Anual d'Estudis Hebraics 8, 43–75.
- Fishman, Joshua A. (1991), *Reversing Language Shift: Theoretical and, Empirical Foundations of Assistance to Threatened Languages*, Bristol, Multilingual Matters.
- GIEC = Institut d'Estudis Catalans (2016), *Gramàtica de la llengua catalana*, Barcelona, Institut d'Estudis Catalans.
- Gumperz, John J. (1964), *Linguistic and Social Interaction in Two Communities*, *American Anthropologist* 66/6, 137–153.
- Haugen, Einar (1968), *The Scandinavian Languages as Cultural Artifacts*, in: Joshua A. Fishman/Charles A. Ferguson/Jyotirindra Das Gupta (edd.), *Language Problems of Developing Nations*, New York, Wiley, 267–284. Collected in: Einar Haugen (1972), *The Ecology of Language*, ed. Anwar S. Dil, Stanford, CA, Stanford University, 265–286.
- IEC (1913) = Institut d'Estudis Catalans, *Normes ortogràfiques*, Barcelona, Institut d'Estudis Catalans.
- IEC (1917) = Institut d'Estudis Catalans, *Diccionari ortogràfic: precedit d'una exposició de l'ortografia catalana segons el sistema de l'I. d'E. C.*, redactat sota la direcció de Pompeu Fabra, Barcelona, Institut d'Estudis Catalans.
- IEC (1990) = Institut d'Estudis Catalans, *Proposta per a un estàndard oral de la llengua catalana*, vol. I: *Fonètica*, ed. Joan A. Argenter, Barcelona, Institut d'Estudis Catalans.
- IEC (1992) = Institut d'Estudis Catalans, *Proposta per a un estàndard oral de la llengua catalana*, vol. II: *Morfologia*, ed. Joan A. Argenter, Barcelona, Institut d'Estudis Catalans.
- IEC (2017) = Institut d'Estudis Catalans, *Ortografia catalana*, Barcelona, Institut d'Estudis Catalans.
- Institut d'Estudis Catalans. Secció Filològica (1990), *Reincorporació al Diccionari General de la Llengua Catalana de les supressions atribuïbles a la censura*, in: *Documents de la Secció Filològica*, vol. I, Barcelona, Institut d'Estudis Catalans, 73–75.
- Lüdtke, Jens (1984), *Katalanisch. Eine einführende Beschreibung*, München, Hueber.
- Milián-Massana, Antoni (ed.) (2012), *Language Law and Legal Challenges in Medium-Sized Language Communities. A Comparative Perspective*, Barcelona, Generalitat de Catalunya, Institut d'Estudis Autonòmics.
- O'Rourke, Bernadette/Pujolar, Joan (2013), *From Native Speakers to "New Speakers" – Problematizing Nativeness in Language Revitalization Contexts*, *Histoire Épistémologie Langage* 35/2, 47–67.
- O'Rourke, Bernadette/Pujolar, Joan/Ramallo, Fernando (2015), *New Speakers of Minority Languages: the Challenging Opportunity – Foreword*, *International Journal of the Sociology of Language* 231, 1–20.

- Pietikäinen, Sari, et al. (2016), *Sociolinguistics from the Periphery. Small Languages in New Circumstances*, Cambridge, Cambridge University Press.
- Polanco, Lluís B. (1984), *La normativa al País Valencià. Problemàtica i perspectives*, in: *Problemàtica de la normativa del català. Actes de les Primeres Jornades d'estudi de la llengua normativa*, 1983, Barcelona, Publicacions Abadia de Montserrat, 107–146.
- Prat de la Riba, Enric (1906), *La nacionalitat catalana*, Barcelona, L'Anuari de la Exportació.
- Prat de la Riba, Enric (1908), *Importància de la llengua dins el concepte de la nacionalitat*, in: *Primer Congrés Internacional de la Llengua Catalana. Barcelona. Octubre de 1906*, Barcelona, Joaquim Horta, 665–669.
- Pujolar, Joan/Puigdevall, Maite (2015), *Linguistic “Mudes”. How to Become a New Speaker in Catalonia*, *International Journal of the Sociology of Language* 231, 167–187.
- Verdaguer, Jacint (1907), *“Càntic dels càntics”, precedit d’“Els jardins de Salomó”*, Barcelona, Tipografia L'Avenç.
- Vila, F. Xavier/Bretxa, Vanessa (2013), *The Analysis of Medium-Sized Language Communities*, in: F. Xavier Vila (ed.), *Survival and Development of Language Communities: Prospects and Challenges*, Bristol, Multilingual Matters, 1–17.

Joan F. Mira

1 Languages, Cultures, Nations: A History of Europe

Abstract: In modern and contemporary history, languages in Europe cannot be dissociated from peoples or nations, and nations and peoples cannot be easily dissociated from languages. This association has certain aspects and effects that are currently affecting the sociolinguistic reality, ideological debates as well as social and political life in the Catalan Countries. In modern times, many peoples or societies have become national communities – and ultimately states – primarily because of their survival as linguistic communities and of the consciousness of unity and distinction gained therefrom. Establishing a standard or an accepted codification of the language is usually a strategic objective in affirming a society’s national culture and, because of its very existence and projection, it may also be a decisive factor in the formation of that society. In the Catalan-language territories throughout the Middle Ages, the language community also exhibited a number of shared cultural traits, but it would be a mistake to identify the particular scope of a language and the framework of a culture in overly general terms; the issue is more complex than it seems. However, sharing a basic written-language model means sharing a body of literature and the same pantheon of renowned writers. Many other things go hand in hand with a national literature, including a sense of assumed common identity among the readers and speakers of that language, of belonging to the same mental space and of shared references to the same “moral territory”.

Keywords: nationality, national ideology, national identity, national language, name of the language, people, spirit, language community, state, nation, literature, literary language, Castilian

1 A people without a language of its own

“A people without a language of its own is only half a nation. A nation should guard its language more than its territories – ‘tis a surer barrier, a more important frontier than fortress or river” (Davis 1914, 173).¹ These are the words of an Irish patriot, cited by Carl D. Buck in a classic article on language and the sense of nationality, published in 1916 in the *American Political Science Review* (1916, 48). Note the date and place: 1916, Ireland, a country in which a true war of independence was unable to restore

1 Parts of this text heavily rely on or have been taken substantially from Mira (2006; 2016).

life – active social life – to a language that was practically extinct. In the same article, Buck (1916, 49) goes on to cite the scholar Mahaffy, who had little sympathy with the cause, who stated: “It seems to be a profound mistake that distinct nationality can only be sustained by a distinct language”. Curiously, a few years later, the president of the Republic of Ireland, Éamon de Valera, said that if he had had to choose between the language and independence, he would have gone for the language. It is not hard to guess why. In any event, the patriot’s and the scholar’s statements seem to express, without further qualification, two sides of an age-old debate. Should references to more authoritative figures from bygone eras be required, I would be remiss not to mention, on the one hand, the claims of an English prelate at the Council of Constance in 1414, at the time of the Great Schism of Western Christianity, a time of confusion and of pre-national affirmations, demanding their own separate representation by invoking “*diversitatem linguarum, quæ maximam & verissimam probant nationem, & ipsius essentiam, iure divino pariter & humano*” (Mansi 1784; 1065); “by difference of language, – which is the chief and surest proof of being a nation, and its very essence, either by divine or human law” (Crowder 1977, 120). It would be difficult to find a more robust declaration on the language-nation identification than that made by these conciliar fathers. On the other hand, at the opposite extreme we find the statement from Antoine Meillet, one of the fathers of sociolinguistics: “Une nation n’est pas liée à tel ou tel soutien matériel, et pas même à la langue. Appartenir à une nation est affaire de sentiment et de volonté”. As a linguist, Meillet was influenced by Ferdinand de Saussure, but in these statements the teachings of Ernest Renan are clear to see (i. e. the nation as a “*plébiscite quotidien*” or ‘daily plebiscite’, an uncertain and fraudulent concept), as is Meillet’s faithful adherence to French national ideology, which, at that moment in history – these statements were published in 1918 – went to great pains to demonstrate that Alsace and Lorraine were part of France because of ‘feeling and will’, ignoring the past or present *diversitas linguarum* as a potential barrier to national unity.

This chapter is not concerned with discovering the true foundation of these concepts or determining whether the 15th-century bishops, early 20th-century linguists or the Irish patriots were right. This is because, among other reasons, on this issue of language and national identity – in the modern and contemporary history of Europe at least – all participants in the debate have tended to act as committed patriots rather than as methodical, dispassionate observers. When I say “all participants”, I mean politicians and governments, practical and theoretical nationalists, military personnel, teachers and lecturers, journalists and the general public. In other words, in reality it is not a question of knowing whether a language, in and of itself, is or is not a defining trait of national identity, but of knowing whether there is some substantial link between this criterion and certain others when it comes to framing the space of a national culture and, ultimately, the space of identities considered national. This is the question that needs to be answered, and it is a question that has certain aspects and effects that are currently having a very obvious impact on the

sociolinguistic reality, the ideological debates and civilian and political life in the Catalan-language territories. Here, as in so many other spaces and countries in Europe, the current language conditions are the outcome of a long history – which can often be traced back to the Middle Ages – of certain ideologies that are not always explicit, and of political power games and border changes or continuities.

In the case of the Catalan language, the establishment and spread of the language was clearly a process that went hand in hand with the territorial expansion of the monarchs of the House of Barcelona, the kings of Aragon, which is synonymous with the protracted process (in parallel with the other Hispanic kingdoms) that we often somewhat inaccurately refer to as the “Reconquest”. In the case of Catalonia, unlike that of Portugal or Castile, this was not a straightforward expansion of existing kingdoms, but instead the creation of new, politically autonomous spaces: the short-lived Kingdom of Majorca, the duration of which was ephemeral, and above all, the Kingdom of Valencia, a political structure that remained unchanged until the 18th century. Thus, when the population of the new kingdom (the new Christian population, of course, which was largely descended from immigrants from Catalonia and who gradually became the majority, and not the pre-existing Muslim Arabic-speaking population, who were finally expelled in 1609) began using a name for the language they spoke, they chose the name of the political territory, that is to say, the “Valencian language”. Though the Catalan resettlers obviously brought the language they spoke in their country of origin with them, at the time – from the 13th century to the first half of the 14th century – the language spoken by these immigrants still did not have its own name. They spoke forms of Old Catalan known as *pla* or *romanç*, and with this new language yet to be called “Catalan”, these immigrants spread throughout the new kingdom. The name of the new kingdom would ultimately provide the name for its inhabitants and the language they spoke: “Valencian” was, therefore, the name for the language of the Valencians. Initially, the name was not presented in opposition to the original “Catalan”, but rather to designate the identity of a new political space, the Kingdom of Valencia, which was institutionally (laws, parliament, currency, etc.) different to Catalonia. In this case, these foundations historically conditioned, and still condition, certain phenomena of association between the politico-institutional territory, the demonym, the language and the name of the language, which continue to form the basis, even now, of much confusion, confrontation and conflict in Valencia. In the case of the Catalan language in the Valencian Country, this very specific “question of names” is not primarily a question of Romance philology or of dialectology, or even a sociolinguistic issue; it is an example or manifestation of a deep-rooted phenomenon, of an explicit or implicit ideology, that is present throughout Europe.

2 Is “speaking” “being”?

The issue that concerns us here, then, is not simply an academic debate about language or about “this” language, but about something else instead. This is because, in the history of Europe, “recognising” the distinct and separate reality of a language – especially if it is a language “of culture” – is actually also a matter of recognising, in a more or less explicit way, that there is a people, country or nation, or set of countries, for whom this language serves as the national language. It is a relationship in which both high-ranking politicians and ordinary people believe, be they naïve or ideologists. They believe in it because, explicitly or otherwise, they believe that speaking is being. They may believe this in a tolerant, polite way, or in a categorical way such as that seen in the official, oral and written instructions circulated throughout the Catalan Countries in the 1940s and 1950s, which, depending on how you look at it, still seem to be circulating: “Si eres español, ¡habla español!” An instruction or demand implying that whoever is must speak, and whoever does not speak is not; or, at least, is not in a way that is as genuine and complete as those who speak. Today still, as always, in their hearts – a law unto itself: “le cœur a ses raisons...” – most Spaniards do not believe that the Catalans, who do not speak like the others, are really Spaniards like the others. The Romantic poet Alfred de Musset expressed this thought very succinctly, stating that “celui-là seul est vraiment français du cœur à l’âme et de la tête aux pieds qui sait, parle et lit la langue française” (apud Monzie 1925, 336–337).

The nation-language ideology is unmistakable in the political thought of the French Revolution, and one of its most distinguished intellectual heroes, Abbé Henri Grégoire, formally proposed the “annihilation” not only of *patois* but also of the languages of minority communities, such as Yiddish or Creole, as a way of “fondre tous les citoyens dans la masse nationale” and of “créer un peuple”. At the National Convention in 1794, Abbé Grégoire presented the famous *Rapport sur la nécessité et les moyens d’anéantir les patois et d’universaliser l’usage de la langue française*, an explicit proposal to “uniformiser le langage d’une grande nation”, a glorious undertaking that no other people had yet fully executed, but which “est digne du peuple français, [...] qui doit être jaloux de consacrer au plus tôt, dans une République une et indivisible, l’usage unique et invariable de la langue de la liberté” (Grégoire 1975, 302). However, there was obviously no *liberté* in it at all. Thus, in the two centuries following the Revolution, the dogma of the Trinity, “La France, les Français, le français”, three concepts and a single substance, would be the immutable doctrine of faith (albeit a secular faith). In 1925, the Minister of Education Anatole de Monzie, in a circular prohibiting the teaching of any *regional language*, presented this lovely idea: “L’École laïque, pas plus que l’Église concordataire, ne saurait arbitrer des parlers concurrents d’une langue française dont le culte jaloux n’aura jamais assez d’autels” (Monzie 1925, 210). A *secular* ideal, with all the cults and altars to an implacable, jealous divinity that does not tolerate rivals. And this is how things have continued in

the early 21st century in gentle France, the birthplace of human rights and freedoms. It is not at all odd, therefore, for any French government to have little sympathy for the formal presence of Catalan, either in its own territory or in the institutions of the European Union; who knows whether an official stateless language (and, according to them, a nationless language, too) might trigger uncomfortable comparisons with the Bretons and the Alsatians, for example? It is worth mentioning in passing that it was in this 1925 ministerial circular that the words of Musset were cited as an authoritative argument in support of the idea that the language and the essence of the French are one and the same thing; a metaphysical and poetic idea that, if expressed by a French (or Spanish, Italian, Hungarian, Swedish or Portuguese) minister, is unquestionable and beautiful, but if it is expressed by a Catalan intellectual or politician, triggers accusations of essentialist nationalism or, indeed, of something much worse. *Ideas*, on this issue in particular, are usually valued positively or negatively depending on which ideological perspective they are seen and considered from. “La sangre de mi espíritu es mi lengua / y mi patria es allí donde resuena / soberano su verbo, que no amengua / su voz por mucho que ambos mundos llene. / Ya Séneca la preludió aun no nacida, / y en su austero latín ella se encierra.” These well-known verses by Miguel de Unamuno, for example, distil – in a few sentences – an entire ideology comprising images, concepts and explicit references connected with language: blood, spirit, homeland, the extent of a powerful space and distinguished antiquity that dates back to Latin... from a Basque author who defines himself as profoundly Spanish. It would be impossible to distil so much in fewer words. And patriotism, however one wishes to define it, including political patriotism, can be perfectly reduced to or distilled in language: “Não tenho sentimento nenhum político ou social” (‘I don’t have any political or social feeling’), said Pessoa. “Tenho, porém, num sentido, um alto sentimento patriótico. Minha pátria é a língua portuguesa” (Pessoa/Soares 2010, vol. 1, 326). (‘I do however have a strong patriotic feeling. My homeland is the Portuguese language’). In any event, it is obvious that if the beliefs and the concepts of the French ideologists, of the Spanish philosopher and of the great Portuguese writer were indeed applied universally, the blurring of limits, boundaries, languages, homelands or nations would be indescribable.

But concepts and metaphors like those we have just mentioned are not just products of Romantic or the most extreme political nationalism. They were already circulating during the Enlightenment, and are a very ancient and reputed resource: ‘Our human language is created, as it were, more for the heart than for reason’, said Herder (1989, 352: “ja gewissermaße [sic] ist unsre menschliche Sprache mehr für das Herz, als für die Vernunft geschaffen”). And it is a well-known fact that a people has, normally exclusively, a more or less pure and powerful, more or less great or indestructible *spirit*. This is obviously a phenomenon that is difficult to subject to empirical verification, but this does not usually take away any of its weight or efficacy. Faith in the existence of such a spirit (*Volksgeist* as the German Enlightened and Romantics call it, a word that is now widely used) serves as part of the basis of the European and

non-European national and nationalist ideologies of the last one-and-a-half to two centuries. From here, it is only a small step to the belief that, for each people ‘their language is their spirit, and their spirit is their language’, as Wilhelm von Humboldt wrote (“ihre Sprache ist ihr Geist und ihr Geist ihre Sprache”, 1963, 414–415; written between 1830 and 1835). Let’s leave aside the habitual stereotypes, among which we cannot fail to find the elegant precision of French, the musicality of Italian or the energy of Spanish. Let’s also leave aside the idea that different languages can be better or worse suited to certain intellectual or aesthetic endeavours. There is the central idea, an explicit, powerful metaphor that peoples have a soul or spirit, and language is the organ that exposes and expresses it. When considered calmly and rationally, the things we come to believe are very surprising. The reality is that those supposed virtues of language as a manifestation of a collective spirit are precisely those that serve to *oppose* (that is to say, to distinguish) one “spirit” from another, one people from another. Could it be that the French spirit is logical and precise, and that the French *therefore* speak in a logical and precise language? And is it possible that the Catalans are a relatively straightforward people, who get down to work... and that their language is *therefore* efficient, with short words, allowing them to say things in far fewer syllables than in Castilian? Saint Vicent Ferrer, a Valencian preacher who was very popular throughout Western Europe, explained this in the early 15th century: “Los castellans són molt parlars: Ferran Ferrandec de los Arcos de los Mayores...” (“Castilians are very talkative: Ferran Ferrandec de los Arcos de los Mayores...”). And, in the 18th century, our language apologists busied themselves with giving demonstrations along similar lines. As did some erudite, Enlightenment or Romantic Valencian “apologists”, for whom the “gentleness” and “sweetness” of Valencian was evident, as opposed to the roughness and coarseness of Catalan. What matters above all is that the generalised belief in the correlation between the language of a people or country and its hypothetical spirit – regardless of the dubious objectiveness or grounds on which it could eventually be based – usually has a real effectiveness for collective consciousness, which is the function of contrast or of opposition: we speak like this *because* we are like that, others speak differently *because* they are different or unlike us.

We should also recall that, by virtue and as a result of such a common belief as *speaking is being*, it is helpful to refer to the French national ideology because it is a perfect expression of the deliberate imposition of this principle of identity. However, it is a principle that can also be viewed from a different perspective; that of historic processes in which *speaking* was, for a very long time, the only and principal way of preserving one’s *being*. It goes without saying that here I am referring to those societies that, in modern times, have become political communities – national communities and ultimately states – primarily because of their survival as linguistic communities and of the consciousness of unity and distinction gained therefrom. We could mention cases like Slovakia, Slovenia, Romania, Macedonia, Estonia or Latvia; those peoples that some have referred to as “peoples without history” (that is to say,

without an autonomous political history in recent centuries), but who were nevertheless peoples *with their own language*. Or Lithuania, which has a glorious dynastic and territorial history to its name, but where the name of the Grand Duchy covered regions that were mostly Ukrainian, Polish or Belarussian, and with Russian or Polish as the cultured and public languages. Thus, in the mid-19th century, Lithuanian was only the language of the peasants in a few poorly defined Baltic counties. However, thanks to this language, the Lithuanian nationality survived when those counties, in the hands of the Russians or the Borussians, were the sole testimony of the former power and glory of the Grand Duchy. In Estonia, the (small) urban population spoke German, Yiddish or Russian, and they only began to think of themselves as a potential nation when the “intellectuals” – that is to say, the “writers”, those who write – discovered that the strange dialect the illiterate rural population spoke was a true language, a sister or cousin of Finnish. A strange coincidence arises, almost without exception from observing this aspect and dimension of European history, especially that of the 19th and 20th centuries; namely, that it is difficult for there to be a “national question” without there also being a “linguistic question”. And vice versa, seldom is there a conflict of languages that does not follow, become or express a conflict of a national nature: of societies that want to rebuild or assert themselves in some way (often *against* another society) as communities of culture; of societies that also want to assert themselves in many cases as political communities. In this field, languages are neither “innocent” nor neutral. They are as neutral as the choices made by writers who, under circumstances like these, decide to create literature in one language or another, in the hegemonic or state language, or in the stateless or powerless language: in Hungarian or in Slovak, in German or in Slovenian, in Russian or in Estonian, in Spanish or in Catalan. In cases like these, at the early stages of what we could call “renaissance” or reconstruction of one’s own national space and of the common consciousness that this entails, writers were, above all, the genuine “creators” of a language that had previously not been recognised as such, and therefore of a space for cultural production in this language.

In any event, there is no example or argument more effective and visible as language (when it is one’s own and distinctive, as it is in most cases) on which to base this continuing and former reality. If a different language *already existed* five, seven or ten centuries ago, it means that a different people *already existed*. And if the language *still* exists in the 19th, 20th or 21st century, it means that this people still exists; the continuity has not been broken, we are the same as we used to be. Maybe not exactly the same as before, but the same nevertheless. The considerable element of fantasy that this belief entails does not take away the slightest bit of effectiveness or mobilising power, but the opposite in fact. Nor does it take away the “moral” strength from the argument that says: we exist and we have a language, we are a “country with a language”, and since we want to be known and recognised as a people or country, we also want our language to be known and recognised. Because, if our language is not recognised, it means that we are not recognised as a country. Presenting oneself to the

outside world in one's own language (the Greek word for "language", *idioma*, means just that: one's own thing) is the other side of the correlation between nation and language. It is precisely when one perceives the danger of a break in continuity with the past, the danger of what one was, or what one believed or imagined one was, ceasing to be, that incipient or developed nationalist movements link the project of national preservation to the linguistic project. *Authenticity* and *unity* – the objectives of every strand of classic nationalism – will also be the objectives of a programme put forward to assert, or regain, the *national* category for one's own language. In this programme, "national" means a relationship between language and society such that the latter considers the former to be unique, effectively identifying it as its own. All national societies in contemporary Europe – all peoples with a language – that have a constituted state or aspirations to achieve some form of state, have had a "programme" of this kind, at least since the late 18th century (as we have seen in the case of France), and more so in the 19th and 20th centuries. In its most developed form, the objectives are always the same: a *unified* language, more or less with the features and contents that we would today call a basic common and accepted standard, against the dangers of disaggregation; a language *purified* of anything that is not authentic, of anything seen as contaminating, foreign, alien, strange, etc.; and a language that is effectively *national* and *one's own*, an official language that is used formally and in public life, which occupies and reoccupies every space that the language that is not one's own – that of another nation – has managed to occupy.

It should be said that this project is generally not a "peaceful" project. By that I mean that it is not usually applied or does not usually work without external or internal resistance. One such source of resistance is the language that has to be displaced: for example, the failed resistance of German in Bohemia, which was completely displaced by the politically and culturally hegemonic language; the resistance of Spanish in the Catalan Countries, the success of which has been notable thus far, with Spanish maintaining a position of clear dominance and supremacy over Catalan in the social, economic and political arenas, an even more intense supremacy in the media field, and an almost lethal effect in the Valencian Country; or the resistance of English in Ireland, the success of which has been almost total and has led to the almost total disappearance of Gaelic in day-to-day life (despite being one of the country's official languages...). There may also be endogenous resistance to a unifying standard: the secular conflict between Landsmål and Riksmål (or Bokmål and Nynorsk) and their derivatives in Norwegian, or between Katharevousa and Demotic Greek in Greece, the regional mistrust of Unified or Standard Basque, the as yet impossible reunification of Galician and Portuguese, with the seemingly definitive prevalence of the Galician-language standard that is much closer to Castilian. Often, when it comes to regaining or developing a language, the use made of it by writers has been crucial; without the early 20th-century writers in Catalonia, for example, the majority of whom quickly accepted the grammar rules proposed by the philologist Pompeu Fabra and the Institute of Catalan Studies (IEC, as abbreviated in Catalan) –

or the Valencian writers who, in 1932, agreed to abide by the so-called “rules of Castelló”, which were basically the same as Fabra’s and the IEC’s – it would have been difficult for contemporary Catalan to attain clear and extensive normalisation as a literary language and, therefore, as a formal language of culture or administration, or as an institutional and national language. This is equally true in the specific case of the Valencian Country, despite the minor conflicts that periodically arise, driven by certain conservative minorities that, often for ideological and political motives, continue to refuse to accept the definition of Valencian speech forms as a variety of the Catalan language, a refusal that, deep down, means subconsciously accepting the belief that the linguistic “Catalanness” of the Valencians would also imply national Catalanness.

In any event, establishing a standard or an accepted codification – when it has yet to be attained – is not only a primary strategic objective in affirming a society’s national culture. Because of its very existence and projection, it may also be a decisive factor in the formation of that society. When, from the 17th century onwards, Slovak Catholic clerics began disseminating religious literature in the country’s language, they were laying the foundations for turning what was previously seen as a Czech dialect into an autonomous literary language, in such a way that, later on (when Slovakia became part of the Kingdom of Hungary, and the Budapest government harshly imposed Hungarian on the region), resistance to the Magyarisation did not give rise to a pan-Bohemian nationalism, as preached by Prague, but instead to an autonomous assertion of the Slovak nation. In the Netherlands, the historic course of autonomy in the late Middle Ages, the independent political assertion of the powerful urban middle classes (the United Provinces were the only territory to remain outside the boundaries of the Holy Roman Empire of the Germanic Nation from the 16th century onwards), led to a variety of Low German – like others – becoming the formal, literary and autonomous standard language. Thus, the opposition between standard Low German and High German became real and nationally effective in the Netherlands, and even became an opposition between Dutch and German. Such opposition had only been seen previously between the “Latin” and Germanic peoples and languages: “Kerstenheit es gedeelt in twee: / die Walsche tonge die es een, / d’andere die Dietsche al geheel” (‘Christianity is divided in two; the Romance language is one part and German as a whole is the other part’), wrote the 14th-century Dutch poet Jan van Boendale. It is a well-known fact that Serbo-Croatian, a language with a compound name, once used a dual writing system: Cyrillic in Orthodox Serbia and Latin in Catholic Croatia. During the supposedly happy times in the state of Yugoslavia made up of more or less reconciled peoples – the times of Josip Broz, Tito – this duality of writing systems did not hinder a growing rapprochement, with the majority of speakers coming to recognise them as a single language. Even the Serbians began to use the Latin writing system more and more often (as could be seen on the streets of Belgrade in the late 1980s). The disintegration of the federation, the bloody conflict and the rekindling of ancient hatreds have led Serbians and Croatians

(and their respective governments, schools, books and newspapers) to assert that they speak different languages, leaving aside the fact that the common name could hardly work as it used to in the newly formed Bosnia. Indeed, since the 1990s, Bosnians – especially Muslim Bosnians, who do not consider themselves Croats or Serbs – have been promoting the autonomy of a “Bosnian” language that, until that time, had essentially been perceived to be Serbo-Croatian. Just remaining in Europe, the list of examples is endless, and they all have a common denominator: the close ties between language and politics, between language and power, between the existence of a “cultured” language and its institutional condition (as a reality or as a project), and between this condition and its internal or external affirmation. The social and political “category” of language, which, as everyone knows, is “superior” to the lowest category – a dialect – is not something that depends on the distinctions of philological science, but instead on the effective dissemination of an autonomous codification; people perceive that they have their own language when they know and accept that there is a *language* model for their *speech*. This does not necessarily have to be a model of concepts of “parole” and “langue”, but rather the existence of a model seen as correct, literary, formal and, if possible, institutional, considered as one’s own and corresponding to one’s own language. This reference then attains a very effective symbolic value, it becomes proof and a demonstration of the language’s autonomy and individuality, and (especially in the case of a language that is, or is considered, exclusive) a mechanism for asserting and strengthening the distinct identity of the country or people that speak it.

Hence, languages, and especially so in Europe, do not only have functional-rational value of an instrument or vehicle for communication. If languages possessed solely this “utilitarian” value, languages that are not as widespread – or are spoken by fewer people – would easily be relinquished in favour of more widely spoken ones... and this is what those who subscribe to this view really mean. One’s “own language”, the *idioma*, has an added value of identity, dignity and representation: – obviously not language in general, but rather *this* particular language – which, when seen as codified and cultured, becomes in some way a symbol of itself, with a value and dignity equal to those of other languages. The perception of this recognised “equal dignity” is without doubt essential to the effective perception of the “particular dignity” of the group, society or country that speaks it, and it is essential for ensuring that those mechanisms of cohesion and adhesion function; it is not easy to adhere or be faithful to something – a language or social group – that is seen as inferior or lower in value. And “value” and dignity are the outcomes not only of knowledge, but also of recognition and, if possible, of some form of prestige. As in the case of Catalan in recent decades, being a language that has been the guest of honour at major international book fairs (in Turin, Guadalajara and Frankfurt), possessing one of the broadest global collections of dual-text Greek and Latin authors, having classical and contemporary literature translated into dozens of other languages, being taught in 150 foreign universities, etc., are accomplishments that

consolidate and strengthen the loyalty of speakers of the language and the perception that the language has dignity and “value”, without which the robustness of a language, under the circumstances of today’s societies, is always at risk of weakening and of dissolving. This is because the nation itself and its own people, the basic identity group to which one belongs, cannot be seen as unworthy of wider recognition or of being seen as equal to other national languages; otherwise it would be under threat of experiencing some of the many forms of alienation or collective schizophrenia, or some of the many symptoms of extinction by dissolution. Nor can one’s own language be seen as inferior, inept, unworthy of recognition, otherwise one would be at risk of abandoning one’s own language out of a sense of hopelessness or coercion. As Meillet (1918, 98) himself stated, “une langue ne subsiste que difficilement et misérablement là où elle n’est pas soutenue par un sentiment national”. This observation can also be applied to the history of modern Europe, from Estonia to Portugal, from Sweden to Greece.

3 On culture and language

It is also necessary to take a glance at the field of concepts and provide another brief review of history. Firstly, it is not necessarily helpful to identify (as is usually done in the Catalan-language countries and all other countries in Europe, generally in good faith) the particular sphere of a language and the framework of a culture. The issue is more complex than it seems, and it is not possible to go into much detail here. Nor can we leave aside some of the more overly simplified assertions in this field or matter. So, I shall make some basic assertions knowing full well they are very much open to debate. The coincidence of a physical space with a culture and a language, whatever meaning we give to the first term, may occur more easily in smaller, relatively simple societies, and under conditions of scant intercommunication. In Classical Greece, for example, where Greek was spoken, we can be sure that the non-linguistic aspects of culture – art, architecture, worship, politics, etc. – were also substantially Greek. That was until the Hellenist and Roman period, when many urban sectors of culture, which could scarcely be described as “Greek”, knew and spoke Koiné Greek, which served as an “international” vehicle for communication, a bit like English does today. In most tribal societies, we also have the same margin of probability: if people speak Nambya in Zimbabwe, we already know how they bury their dead; if people speak Sindebele, we already know how they build their huts and what kind of feast they hold on wedding days. But we cannot be quite so sure when spaces are much bigger and complex, as was the case for the Roman empire, for example: Latin was the cultured, urban and official language throughout the whole empire, and the same “cultured” architecture and administrative structure could also be found in all of the empire’s provinces. But between Egypt and Britain, for example, the ethnic differences and distances of all kinds were so visible that it would make little sense to assert that the Egyptians and the

Britons shared the same culture. We could speak in similar terms about the culture and cultures of Mediaeval Europe, united by the Church, by Latin, by Romanesque or Gothic architecture, by the feudal system and by so many other things, but which gradually became separated by different languages that gave rise to diverse literatures and to communicative spaces that increasingly came to coincide with the territory of the language. In the 15th century, Catalan culture was, on the one hand, a regional expression of Western European Christian culture, and as such was no different from Italian, Castilian or French. However, it had developed a set of traits that emerged particularly in the territories where Catalan was spoken, such as certain characteristic aspects of trade relations, of institutions of government or certain variants of the style of architecture that centuries later would be called Gothic. Moreover, an abundant body of Catalan literature had begun to develop some 200 years previously, which, leaving aside the content and specific forms evident therein, stood apart from the other national literatures because of the language in which it was expressed.

It is clear that, in the 14th and 15th centuries, nations still did not exist in Europe with the political structure, precise boundaries and intensity of consciousness that they would acquire in subsequent centuries. However, there were “peoples”, pre-nations or quasi-national societies, including the Portuguese, French, Castilians, English or Catalans, for example, who were “peoples” in a much more defined way than the Germans or the Italians. While others, such as the “Spaniards”, “Belgians”, “Swiss” or “Austrians” simply did not exist as peoples or nations, nor did many of the Eastern European societies, which were still not characterised as nations in any way by modern standards. In other words, they did not have a written culture in their own language, an institutional definition or an urban network. In our case, while the Kingdom of Valencia and the Principality of Catalonia – with the Kingdom of Majorca – strictly speaking did not form a single state, they did in fact find themselves within the same sovereign space, they shared parallel public institutions and, above all, they shared the same urban-based cultural complex and the same formal, cultured and written language. Thus, alongside texts expressing a consciousness of unity, there was, above all, the *reality* of this pre-national ensemble of Catalan-speaking people, of this ethnocultural unity, and, in short, of this common “nationality”, culturally at least, in terms comparable to the clearest and most defined nationalities in Western Europe. We can make this assertion without retrospective recklessness and without forcing the facts, more than anything because the Valencians shared a very robust set of cohesive factors with their neighbours to the north and on the islands, among which of particular note are:

- a) The origin of the majority of the Valencian population, not only in the first century of the Kingdom’s founding, but also from constant immigration throughout the following two centuries. The Valencians were fully conscious of this community “lineage” (it had never been denied, and it had repeatedly been affirmed), even when they had wanted to distance or distinguish themselves from their country of origin; even Eiximenis, when highlighting the existence of a “Valencian people”

distinct from the “Catalan people”, did so after reminding us that the former “[és] vengut e eixit per la major partida de Catalunya” (‘mostly came and originated from Catalonia’). It was therefore a matter of “political” personality, and not of “ethnic” difference of origin or lineage. This awareness of a linguistic community and of shared origin, neither contentious nor combative, was maintained throughout the following centuries, before being taken up again by a sector of the late 19th-century and early 20th-century revivalist movement known as the *Re-naixença*. In the 18th century, the Valencian bishop Josep Climent recalled that “casi todos los valencianos somos catalanes en el origen, y con corta diferencia son unas mismas las costumbres y una misma la lengua”.

- b) The same political, legal and trading culture: equivalent or very similar forms of municipal government and territorial government (town and city *customs*, juries, viceroys and governors, autonomous government or Generalitat, etc.), or common institutions in commercial life, such as company and shipping contracts, exchanges or consulates. As well as this, there existed a common space for trade relations, and a certain “Catalan commercial bloc” that shared facilities and maritime routes, consuls and representatives, codes and courts, etc., throughout the Mediterranean and even beyond.
- c) The same language with very little dialectal variety (so little that it is nearly impossible to place the geographical origin of a 14th- or 15th-century text by language alone, or to know whether the person who wrote it was Valencian or not). Catalan was uniform as a popular spoken language, and perfectly stable and established as a written, formal, official and cultured language. It was a language of trade, documents, curia, courts, public registers, treaties, contracts, administration, parliaments, general and local government, etc. Catalan was also a language of books, that is to say, the language of a common literature as mature and “national” as the Castilian or French literatures, for example, and more so than many others in Western Europe. This also signifies the existence of a group of elite readers, aristocrats or members of the mainly urban middle classes, clergymen or professionals across the various regions who were all “consumers” of the same literature, followers of the same fashions and trends, and part of the same “cultural network”.
- d) A cultural community that also includes an ideological sphere that is reflected in the legal thinking of the University of Lleida and of Raimon de Penyafort, in the work by Arnau de Vilanova and *arnaldisme*, or in Ramon Llull and the Lullian tradition. This cultural community includes the spread and circulation of common forms of architecture and other forms of art. Art historians are in agreement over the existence of a specific “Catalan Gothic” style that can be observed right up until the 16th century. This was a style of construction common throughout the territory (already observable, in part, in late 13th-century Valencian churches), and is highly visible in the tendency towards building more compact churches and in the unitary sense of the interior space. It is also visible in shipyards,

merchants' exchanges, in the structure of urban palaces, and in octagonal-foot-print towers.

- e) A common religious and ecclesiastical space: The Diocese of Tortosa extended to the southern border of what is now Castelló Province, and the majority of the remaining territory formed part of the large Diocese of Valencia, a suffragan of the Archdiocese of Tarragona. In the first two centuries of the new kingdom, the bishops of Valencia were Catalan and most of the diocesan clergy and members of religious orders were also of Catalan origin and Catalan-speakers. Likewise, the big Valencian monasteries were subsidiaries of the respective principal religious houses to the north of the Ebre River (Benifassà of Poblet, Valldigna of Santes Creus...). This common linguistic area was the preferred space for the circulation of mendicant and preaching friars – Franciscans, Dominicans, Mercedarians etc. – who had a powerful influence over the people of the towns and cities, in terms of not only education and charity, but also relations with circles of influence and power. Among the most illustrious figures in Catalan religious life, we should mention Francesc Eiximenis, a Girona-born Franciscan living in Valencia and the “ideologist” of the urban middle classes; the Mercedarian Gilabert Jofrè, founder of the “Hospital dels Folls” in Valencia, and commander of his order in Lleida and Perpinyà; and the Dominican Vicent Ferrer, born in Valencia to Girona-born parents, a student in Barcelona and Lleida, and advisor and preacher throughout the Catalan lands and beyond the Pyrenees.

This is not the place to go into greater detail about these factors or certain others, such as those of a more “anthropological” nature (agrarian landscapes, labour and social practices, deep-rooted cultural attitudes, etc.), but it seems reasonable to conclude that the resulting sum and combination of these factors makes it possible to assert the existence perhaps not of a common “national” sphere – in the sense that the word “nation” can be applied to late mediaeval society – but certainly of a community of “people”, of language and different levels and aspects of culture.

In any event, Mediaeval Europe was made up of open spaces, of relatively unstable frontiers and of peoples who did not rigidly form their own states with precise administrative borders. In the modern world, however, these conditions gradually changed, for better or for worse; territories became enclosed by rigorously controlled borders, languages – though not all, of course – became instruments and expressions of states, and some states spread beyond their original limits. Between the 16th and 20th centuries, though we may no longer remember, Europe was primarily made up of expansive monarchies and empires, both internal and external. And, by some means or another, a sovereign language, invariably the language of the sovereign or of the people who identified with it in the most direct manner, would manage to spread throughout each imperial space, e.g. Turkish, Russian, German, French, Spanish or English. However, it would be very difficult to speak of a Turkish culture that stretched from Anatolia to the Danube, a Russian culture from

Georgia to Finland, or of a Spanish culture from Barcelona to Chile. This is not a suitable place to assess the multitude of theories on the topic. But it is an appropriate place to briefly introduce some comparisons that may be useful in order to situate the Catalan language and culture – that is to say, the language *or* the culture – within their context, within their own space and, above all, in relation to the territorial, historic and even ideological and political space of the Spanish language and/or culture.

For example, in the 18th century, in what sense was it possible to speak of a Romanian or Macedonian culture when, in the regions that were subsequently called Macedonia or Romania, the written languages were not Macedonian or Romanian, but instead Church Slavonic, German or Transylvanian Hungarian, the Greek of the traders or the Turkish of the imperial administration? It might perhaps be possible to speak of Romanian or Macedonian culture, but then it would only be an “anthropological” culture (an imprecise expression), a popular culture without a “cultured” dimension and certainly *not* national cultures. And, at the other extreme, what do the cultural expressions of post-colonial Mexico, of Peru and of Argentina have in common apart from the fact that they share the colonial past of the same empire and then went on to share the same written language? What I mean is this: What sense does it make to speak of a “Hispano-American culture” if the Inca, Quechua, Guarani, Maya and Aztec past – and present – are to be included too? It makes little sense, though it does make some sense: above all, it makes some sense in the part of culture that we call literature. And we could say the same thing about German culture: does it, or does it not, include Austria and Switzerland *in the same sense* as Bavaria or Pomerania? And is French culture *equally* French in Quebec and in Martinique, Belgium, Haiti and the Republic and Canton of Geneva as it is in Paris or Marseille? The answer is not straightforward. In the case that concerns us, it is precisely the inclusion of the Catalan space in a more extensive power structure (the Spanish monarchy, which is primarily a Castilian Monarchy) which, from the 16th century, led to the political subordination and the marginalisation and weakness of one’s own language due to the dominance of Castilian, which also meant a weakening of the production of one’s own models of cultural expression and the gradual lack of communication between the urban elites of Catalonia and Valencia. These disastrous effects (subordination, marginalisation, weakness...) became even more intense throughout the 18th century due to the politics and ideology of the new Bourbon dynasty which called for the annihilation of the representative and governmental institutions of the Principality of Catalonia and the Kingdom of Valencia, the administrative and legal incorporation to the Kingdom of Castile, and the imposition of Castilian as the sole official language in all fields of public life, accompanied by repression and the explicit banning of the formal use of the Catalan language. The effects on the vitality of Catalan culture as a “written culture” were devastating.

At any rate, clarifying when and where, and under what conditions one could speak of *a* culture as being coincident with *a* language would take too long and be

overly complicated, and if we wanted to contribute something more than broad generalisations, we would have to specify in what sense we are talking about “culture”, and in what areas there are coincidences or not. Let’s forget about this, then, and talk about literature.

4 Reading and writing: literature

Let’s talk about literature, then, but not without failing to mention a few very obvious and indispensable finer points. Firstly, for the purposes of the issue that concerns us, we shall take “literature” to mean the body of written texts that may not be strictly instrumental, but do have a minimally aesthetic intention. A definition like any other, possibly useless, which simply intends to consider *literature* – regardless of its high or low aesthetic quality or of its cultured or “vulgar” nature – texts in the style of mediaeval sermons by Saint Vicent Ferrer or the verses published in the programmes of popular festivals, but not the telephone directory or the instructions for a washing machine. Secondly, it is true to say that it may make some sense, and perhaps more so than geographically and politico-administratively, to speak of Bolivian or Ecuadorian literature, Belgian or Swiss literature, and Andalusian or Valencian literature. It is always possible to find something in common within a specific territory, whether it be the topics about which authors write (in one or more languages), the tone given by their historical or geographical framework, or the particular colour or flavour of the language. But ultimately, having made all the distinctions you want to make, and having accepted all the territorial or any other kind of variants, *a* literature is *a* literary language, i.e. English, Latin, German, Spanish or Catalan literature. And it means that all the speakers of the same language, with a certain level of literacy and formal education, can *read* the same “literary products” as their *own* products and *not* as “external” products, to which one only has access through knowledge of *another* language (and, in any event, as a product that is not one’s own, but is instead incorporated into one’s own language through translation). And then, to read a literature as one’s own, neither the topics, the location, the vocabulary nor the linguistic variant are important; what matters is the common literary language, as such and only as such, defining a space that separates the external from the internal. Like the members of the Academy of Motion Picture Arts and Sciences in Hollywood, who give the Oscar first to a film in English and then to a foreign-language film. Like the Australian reader who does not consider Shakespeare a “foreign” author. Or the Spanish (from Spain) reader – scholar, lecturer, minister or monarch – who, every year, considers the *Premio de Literatura en Lengua Castellana Miguel de Cervantes* (Miguel de Cervantes Prize) a grand prize for *our literature*, or for *Castilian* or *Spanish* literature. Indeed, by statute, which is tantamount to saying “legally”, the Miguel de Cervantes Prize cannot be awarded to Catalan-language writers (they are not *ours*: they do not belong to “Spanish literature”, they are outsiders, and from this perspec-

tive they are *foreign*), though it can obviously be awarded to writers from Mexico or Peru: they are “insiders”, *ours*, and Spanish.

In Veracruz, Lima and Alcalá de Henares, it is not certain, at any rate, whether the bulk of the inhabitants actually share the same *culture* that transcends oceans and mountains, or even whether the language of the respective popular neighbourhoods is mutually understandable, but what is certain is that people who are more or less “cultured” (those who have several years of schooling, newspaper readers...) only consider *theirs* a literary language: just one, and the same one. That’s why it is so important to bear in mind that each written language corresponds to a *literature*, and only one; because sharing a literature as one’s own also means, and forgive the apparent redundancy, sharing the same written language, the same writers and the same basic language model. And, it goes without saying that many more things come with a shared language: a sense of assumed common identity (of *being*, in some vague or clear way, the same thing that the readers and speakers of that language are), of belonging to the same mental space, and of shared references to the same “moral territory”. To be specific: for reasons and circumstances that, above all, depend on the region’s incorporation into the substantially Castilian monarchy, the few people in the Catalan-language countries who had a certain level of schooling or formal education – primarily the nobility and urban aristocracy – started writing and, above all, reading in Castilian from the 16th century. This does not simply mean that they entered a “Spanish market” (that of touring theatre companies, for example, or of Castilian-language publishers, many of whom set up business in Barcelona or Valencia), but also that they entered a “Spanish space” – Castilian-Spanish, to be more precise – of identity; of an identity that was becoming more “national”, or that was being presented and assumed as such. What I mean is that reading and writing in Castilian progressively meant, and more so as more people became able to read and write, assuming as *their own* the authors of Castilian literature, “learning” (primarily at school) and therefore thinking that Cervantes or Quevedo, and later Moratín or Zorrilla, were great writers that were “ours” – a perception reinforced, above all from the 19th century onwards, by the names of a whole host of squares and streets dedicated to them by municipal authorities in Catalonia, Majorca and the Valencian Country. And, from that point onwards, with the same accelerated progression, we came to consider Pelayo, El Cid Campeador, Hernán Cortés and Pizarro as “ours” and our own: we could go back and take another look at names on the urban street maps. And simultaneously, and inevitably, the perception of a *written culture* in the Catalan language – in Catalan or Valencian – as a culture in its own right gradually dissolved; a culture with its own classics, its own famous names, its own models, its own national literature and everything else associated with it. I am not explaining anything that is not a well-known fact; not writing and, above all, *not reading* turned into not knowing and not feeling. And losing one’s own writers as a point of reference was the first step to losing one’s own country.

If there was anything that the men of the *Renaixença* – some, not all! – had a clear vision of from the very beginning, it was precisely the idea that Joaquim Rubió i Ors expressed in 1841 in the prologue to the verses that he published under the nom de plume *Lo Gaiter del Llobregat*: “Catalunya pot aspirar encara a la independència, no a la política, puix pesa molt poc en comparació de les demés nacions..., però sí a la literària.” (‘Catalonia can still aspire to independence, not to political independence, since it has very little weight in comparison to other nations..., but instead to literary independence’). *Literary independence* indeed; more than a century and a half later, this idea appears highly revolutionary for its time and is even more “novel” and current than it seems. A certain mental and moral independence would ensue, some cultural independence projects and other concepts and programmes for overcoming various forms of dependence, including – almost inevitably – political dependence. As in so many other comparable cases in Europe throughout the 19th century (Finns, Lithuanians, Czechs, Slovaks, Slovenians, Croatians, Hungarians, Romanians and others), it seems that it was necessary to start by claiming a literature and then, or simultaneously, to turn the language into a “conscious model of language” suited to formal uses in education and, above all, in every field of public life. This historical process unfolded in Catalonia in the second half of the 19th century and first third of the 20th century (until the Spanish Civil War, of course), following more or less the same stages and patterns as the other languages and countries listed above.

5 A corollary or conclusion

Hardly anything affecting the state and condition of the Catalan language or anything that might decide its future can be understood – in relation to the language issue and so many others – if we do not at least understand the contemporary evolution of the ideas and facts. For example, if we do not understand that, a century and a half ago, Catalan was one of the many more or less important languages in Europe that, on the one hand, seemed to be condemned to extinction and, on the other, were already being claimed as a language of a people, country or nation that in some way aspired to an effective form of political existence. If the historical comparison is as significant as it is uncomfortable, it is because now, in the second decade of the 21st century, most of those languages (many with fewer speakers and “less history” than Catalan) are effectively languages of a state or a nation, and nobody questions their category or official status either in the United Nations or the bodies of the European Union, whereas Catalan is still denied its place and its rights in this field and in several others. Maybe it would not matter if everything were an accessory matter of formal recognition, but it does matter a lot because, neither two centuries ago, one century ago nor now, languages in Europe could not and cannot be dissociated from peoples or nations, and nations and peoples cannot be dissociated from languages, and this is the case in the history of the social perception, tantamount to the “political” percep-

tion, of one's own language in the Catalan territories and in any other space or territory in Europe.

6 Bibliography

- Buck, Carl Darling (1916), *Language and the Sentiment of Nationality*, The American Political Science Review 10, 44–69.
- Crowder, Christopher Michael Dennis (1977), *Unity, Heresy and Reform, 1378–1460. The Conciliar Response to the Great Schism*, London, Arnold.
- Davis, Thomas Osborne (1914), *Our National Language*, in: Thomas Osborne Davis, *Selections From His Prose and Poetry*, London/Dublin/Belfast, Gresham Publishing Company, 172–179 (Reprint 1982).
- Grégoire, Abbé (1975), *Rapport sur la nécessité et les moyens d'anéantir les patois et d'universaliser l'usage de la langue française*, in: Michel de Certeau/Dominique Julia/Jacques Revel, *Une politique de la langue. La Révolution française et les patois*, Paris, Gallimard, 300–317.
- Herder, Johann Gottfried von (1989), *Ideen zur Geschichte der Philosophie der Menschheit*, Frankfurt am Main, Deutscher Klassiker Verlag.
- Humboldt, Wilhelm von (1963), *Ueber die Verschiedenheit des menschlichen Sprachbaues und ihren Einfluss auf die geistige Entwicklung des Menschengeschlechts*, in: Wilhelm von Humboldt, *Werke in fünf Bänden*, vol. 3: *Schriften zur Sprachphilosophie*, Darmstadt, Wissenschaftliche Buchgesellschaft, 368–756.
- Mansi, Joannes Dominicus (?1784), *Protestationes Anglicanorum*, in: Joannes Dominicus Mansi, *Sacrorum conciliorum nova, et amplissima collectio [...]*, vol. 27, Venetiis, Apud Antonium Zatta, 1058–1070.
- Meillet, Antoine (1918), *Les langues dans l'Europe nouvelle*, Paris, Payot.
- Mira, Joan F. (2006), *A History of Europe. Cultures, Languages, Nations*, Transfer: Journal of Contemporary Culture 1, 6–15.
- Mira, Joan F. (2016), *La nació dels valencians*, Barcelona, Proa.
- Monzie, Anatole de (1925), *Circulaire du 14 août 1925 relative aux idiomes locaux*, Bulletin départemental de l'Éducation nationale. Inspection académique.
- Pessoa, Fernando/Soares, Bernardo (2010), *Livro do Desasocego*, ed. Jerónimo Pizarro, Lisboa, 2 vol., Imprensa Nacional-Casa da Moeda.
- Renan, Ernest (?1922), *Qu'est-ce qu'une nation? Conférence faite en Sorbonne, le 11 mars 1882*, in: Ernest Renan, *Discours et conférences*, Paris, Calmann-Lévy, 277–310.
- Rubió i Ors, Joaquim (1841), *Lo Gayté del Llobregat: poesias de don Joaquim Rubió y Ors*, Barcelona, en la estampa de Josep Rubió.

Joan Julià-Muné

2 History of Catalan Linguistics

Abstract: This contribution analyses the stages and main figures, together with their works, in the history of Catalan linguistics over a century and a half (1858–2018). I begin by detailing how the language was first studied as a relevant Romance language in late 19th-century Catalonia and soon afterwards in the other Catalan-speaking regions. This chapter also highlights the most successful and, therefore, the most prolific periods of Catalan linguistics, namely the first and last quarters of the 20th century, including up to the present day. Whenever the Catalan-speaking countries have enjoyed a certain amount of home rule in peacetime, Catalan language studies have flourished. This chapter presents some of the key works at every level of linguistic analysis and also aims to introduce the ever-increasing number of projects currently being undertaken in many of these areas.

Keywords: Catalan language studies, pronunciation, orthography, grammar, dictionaries

1 Introduction: a century and a half of Catalan linguistics (1858–2018)

In order to analyse and critically evaluate scientific studies, including those about language, as well as to objectively interpret scientific thought, historians of science must display both competence in the given field of study and lack of emotional and personal involvement. This is a question that was dealt with succinctly by Henry M. Hoenigswald: “The historiography of any discipline has its well-known and obvious twofold attraction and twofold challenge: It calls for competence in the history of scholarship and science, and it also calls for a very special kind of competence in the subject field – the ability not only to contribute to it, but to see it with detachment as well. The degree to which this double requirement has been filled must vary greatly across the *globus intellectualis*” (1986, 172).

In analysing the stages and main figures, together with their works, in the history of Catalan linguistics over approximately a century and a half (1858–2018), a statement made by Professor Robert H. Robins springs to mind: “Every scholar is an individual, and ‘schools’ and ‘periods’ are abstractions doing doubtful justice to the work and the workers actually comprised in them” (1967, 209). However, a survey like this is aimed at explaining when and how Catalan first became the subject of academic study as a significant Romance language in the late 19th-century Catalonia up to the present day across all Catalan-speaking regions, mainly in Spain, but also in

<https://doi.org/10.1515/9783110450408-003>

three other states: Andorra, France and Italy. In attempting to highlight the various achievements of Catalan linguistics, we cannot help but mention its most successful and prolific periods, namely the first and last quarters of the 20th century, also extending into the early 21st century. In general, whenever the Catalan-speaking countries have enjoyed a certain amount of home rule in peacetime, Catalan language studies have flourished. In this chapter, the key works from every level of linguistic analysis will be presented along with the ever-increasing number of projects currently being undertaken in these areas.

To begin with, even the most renowned forerunners of Catalan linguistics, as we shall see below, owe their pioneering work in part to their own precursors, who include: Rafael Martí de Vicià (1502–1582; *Alabanzas de las lenguas hebrea, griega, latina, castellana y valenciana*, 1877), Antoni de Bastero (1675–1737; *Història de la llengua catalana*, Feliu 1997; 2000), Josep Ullastra (1690–1762; *Grammatica cathalana*, 1743), Josep Pau Ballot i Torres (1747–1821; *Gramatica y apologia de la llengua cathalana*, 1815), Joan Petit i Aguilar (1752–1829; *Gramàtica catalana*, 1796–1829), Antoni Febrer i Cardona (1761–1841; *Obres gramaticals*, 2004, 2017), Pau Cardellach (1814–1879; *Gramàtica catalana*, 1840) and Antoni de Bofarull (1821–1892; *Estudios, sistema gramatical y crestomatía de la lengua catalana*, 1864).

In the mid-19th century, when descriptive linguistics was still in an embryonic stage within the field of historical and comparative linguistics, Manuel Milà i Fontanals (1818–1884) introduced Romance linguistics in somewhat brisk terms: “[...] las lenguas romances provienen de un latín mal hablado y peor pronunciado, modificado con el tiempo [por] efecto de causas diversas, y que [fue] admitiendo algunos elementos extraños, más o menos considerables, pero no esenciales” (1893, 114).

2 The forerunners: Catalan linguistics in the late 19th century and the *Avens/Avenç* “campaign”

We can thank Milà i Fontanals, as the first Catalan Romance linguist in the 19th century, for the first scientific description of Catalan (1875). Out of all of Milà’s work, his remarkable contribution to dialectology must be particularly highlighted: the division of Catalan into Eastern and Western geolects, based on the law of neutralisation of the unstressed vowels *a*, *e* and *o* in Eastern Catalonia; as well as the definitive establishment of the use of the grave and acute accents to differentiate, respectively, between open and close vowels. For this reason, Milà can be considered a worthy precursor of Catalan linguistics.

Born half a century later, Pompeu Fabra (1868–1948) would become the most prominent grammarian and linguist in the Catalan-speaking regions (Solà 1987a; Segarra 1998; Ginebra/Solà 2007). In fact, he was the creator and main instigator of the spelling, grammatical and lexical standardisation of Catalan. A complete edition

of his collected works was published recently (2005–2013). During the final decade of the 19th century, the first two grammars of Fabra's were published. They were both descriptive and innovative compared to his previous grammatical works: *Ensayo de gramática de catalán moderno* (1891) and *Contribució a la gramàtica de la llengua catalana* (1898). He also published the first scientific work on Catalan to be fully integrated in the new discipline of Neogrammarian historical linguistics (*Étude de phonologie catalane (catalan oriental)*, 1897).

In the Balearic Islands Tomàs Forteza (1838–1898) produced a grammar using a historical-comparative methodology – containing just the chapters on phonology and morphology –, but unfortunately it did not see the light of day until it was edited by Antoni M. Alcover (see 3.1.2) with an extensive and informative preface in 1915, three years after Fabra's descriptive grammar par excellence (1912) was published.

In the field of lexicography, alongside Pere Labèrnia's legacy (*Diccionari de la llengua catalana ab la correspondència castellana y llatina*, 1839–1840), it is worth mentioning the Majorcan Marià Aguiló (1825–1897) and the Barcelonan Josep Balari (1844–1904), whose works were partially published posthumously in the first third of the 20th century. Aguiló meticulously collected a large amount of vocabulary, onomastic data, phraseology and sayings, from both printed and oral sources. The materials were edited by Fabra and Manuel de Montoliu (1877–1961) at the Institut d'Estudis Catalans and published under the title of *Diccionari Aguiló* (1915–1934), which has become an essential work in Catalan lexicography of the 20th century. The Hellenist Balari also made a remarkable contribution to the collection of lexicographical materials that was to be published by Manuel de Montoliu (see 3.2.2).

However, it was the journal *L'Avenç* and its "linguistic campaign" carried out during the period 1890–1892 that would pave the way for the future Catalan language reform under the guidance of Fabra, together with Jaume Massó and Joaquim Casas-Carbó. This was a matter of unifying, cleansing and modernising the language as a reaction to the sort of restraints and narrow-mindedness that had previously dominated the landscape. They proposed ennobling the spoken language so that it could become a language of culture, free from the influence of Spanish and based on the north-eastern variety of Catalan, since this was the most important variety from a demographical and economic point of view, and could also be considered the most differentiated from Spanish. The activist group itself clearly set out its specific aims whilst at the same time keeping away from the historical-comparative approach that was in vogue at the time: 'We prefer to study the spoken language: our research is aimed at formulating its grammar. [...] But our historical investigations on a particular form of Old Catalan grammar should not pretend to be the grammar of modern literary Catalan' (Massó i Torrents/Casas-Carbó/Fabra 1891, cited from Lamuela/Murgades 1984, 163).¹

¹ All quotes in Catalan are translated by the author of this chapter.

Fabra was already announcing the most significant mission that he would undertake during the following century, as Ferrater states:

‘Fabra belongs to a generation of great grammarians of the living European languages [...] who made an accurate intellectual decision to restructure the concepts of the traditional descriptive grammar, and to apply them adequately to Indo-European languages that could be considered ‘old’ – the current ones. Their passionate aim was to describe all languages on their own merits, by struggling against the aprioristic acceptance of the system of the logicist Graeco-Latin grammar, with which young boys are still tortured at school’ (Ferrater 1981, 4).

Even the great European works of the time on Romance linguistics took Catalan into account. One such work was the second edition of the manual of Romance linguistics edited by Gustav Gröber (1904–1906), in which Joseph Sarrailh (1904) expands on the contributions made by his teacher Alfred Morel-Fatio (1888) to the first edition of the work. It was such a relevant work that both Fabra (1907) and Alcover (1909) wrote detailed reviews that became an indispensable accompaniment to the original.

3 Linguistics in the new century thrives under Catalan self-rule (1901–1925)

The President of the Barcelona Provincial Council (Diputació de Barcelona), Enric Prat de la Riba (1870–1917) published his influential work *La nacionalitat catalana* in 1906. This was a crucial year for Catalan linguistics, since the First International Congress of the Catalan Language was held in Barcelona under the joint leadership of Prat and the Majorcan linguist Antoni M. Alcover (1862–1932), who was greatly influenced by the German linguist Bernhard Schädcl (1878–1926) (see 3.2.1). At the time, Prat was de facto president of Catalonia, only becoming de jure president in 1914 when he founded the Mancomunitat de Catalunya, which in time would become the Generalitat de Catalunya, that is, the autonomous government of Catalonia, which was restored in the early 1930s under Spain’s Republican Government. Prat was also the editor of the conservative and nationalist newspaper which, as its title suggests, served as the voice of Catalonia: *La Veu de Catalunya*.

Catalonian self-rule was, therefore, as far as political and cultural projects were concerned, an excellent springboard to achieve the President’s goals, above all thanks to the enormous contribution by Eugeni d’Ors’s daily column (*glosa*, thus *Glosari* as a whole), which was published in *La Veu* from the very first day of 1906. Ors is considered the ‘greater verbaliser of the new century’s spirit (*Noucentisme*)’ (Murgades 1976, 45) due to his philosophical corpus. Prat, together with Alcover and Ors, among several others, helped to promote and found the Catalan Academy, known as the Institut d’Estudis Catalans (IEC) the following year (1907). Four years later a section for the Catalan language was created, the Secció Filològica (‘Philological Section’). At that

point Prat de la Riba appointed Antoni M. Alcover as head of the section and asked Pompeu Fabra to lead its main departments known as the *Oficines Lexicogràfiques* ('Lexicography Offices') to make lexicographical research a main priority. In fact, the departments covered a great deal of work in dialectology, experimental phonetics and grammar, as we will see below (3.2.2). From this basis the Catalan language reform was developed, especially during the first third of the 20th century, under the leadership of the IEC and the engineer (by training, see note 2) and linguist (by devotion) Pompeu Fabra. By the early 1930s, Fabra had almost completed the codification of written Catalan with his *Diccionari general de la llengua catalana* (1932b).

This boom in linguistics in the 20th century lasted throughout the first quarter of the century, until the dissolution of the Mancomunitat (1925). This period can be divided into two distinct stages from the point of view of linguistic studies. The first corresponds to the projects of external promotion (1901–1912), while the second comprises the projects mostly channelled through institutional frameworks, specifically through the *Oficines Lexicogràfiques* of the newly-born *Secció Filològica* at the IEC and the Chair of Catalan Language of the so-called Literary University of Barcelona (Universitat Literària), both under Fabra's supervision (1913–1925).

3.1 Promoters of language studies and personal projects (1901–1912/1925)

Alcover's projects, most of them undertaken in Palma de Mallorca – in the fields of research, publishing and organizing conferences – and those authored by Fabra in Bilbao² – in the areas of descriptive linguistics and grammatical studies – mark the start of the new century and of the first great period of Catalan linguistics.

3.1.1 Antoni M. Alcover's projects: the *Lletra de convit* (1901), the *Bolletí del Diccionari de la Llengua Catalana* (1901–1926/1937) and the First International Congress of the Catalan Language (Barcelona, 1906)

Alcover was one of the leading figures with regard to the recovery of the language at the beginning of the 20th century (Moll 1981; 1983a). He introduced his ambitious first research project on the 'Proposal of the thought', at the beginning of his Invita-

2 Pompeu Fabra held the chair in chemistry in the Basque city between 1902 and 1912, when Prat de la Riba called him, via a close friend of Fabra's Puig i Cadafalch (Prat's future successor in the presidency of the Mancomunitat in 1917), to join the *Secció Filològica* of the IEC. Therefore, whilst contributing extensively to Catalan descriptive linguistics and as a relevant and active participant in the First International Congress of Catalan (Barcelona, 1906), he was also still lecturing in engineering outside of Catalonia.

tion Letter (*Lletra de convit*, 1901), a 48-page booklet – divided into 16 sections and 583 subsections – inviting a great number of cultured people, mainly clerics, in the Catalan speaking regions, to contribute to the project by sending in ad hoc index cards to help him collect written and oral language data for the purposes of the Catalan dictionary that his admired teacher and learned lexicographer Marià Aguiló had not been able to complete. He kept the project alive and published continuous reminders in the journal he founded the first year of the new century: the *Bolletí del Diccionari*, a publication that was to be the diffusion and coordination organ for the dictionary. This is how Alcover introduced his aims in his Letter: '[...] our beloved language [...], in order to be written and spoken properly, needs its own grammar and its dictionary. The almost completed grammar has been left to us, partially printed, by our eminent linguist and most refined poet Tomàs Forteza [...]. However, the dictionary is far from complete' (Alcover 1901, 5).

Alcover organised and presided over the first international conference on the Catalan language (*Primer Congrés Internacional de la Llengua Catalana*, Barcelona, 1906). On 13th October, 1906, Prat de la Riba praised the conference by writing an article in his newspaper *La Veu de Catalunya* entitled *El Congrés d'avuy* ('Today's Congress'). He described it as the 'Catalanist Congress for the language'. There was a massive attendance of around 3,000 congressmen, most of them with little scientific background in the linguistic or historical, literary and socio-legal fields, but who attended out of an enthusiastic sense of patriotism rather than real scientific interest, as Prat alluded to in his article:

'We act instinctively as if we were Germans: we do not keep science separate from the fatherland. [...] Patriotic devotion has given birth to scientific curiosity. All creators of modern Catalan culture have been passionate patriots, Catalans being deeply in love with their land. So, currently it is becoming quite natural to confuse a patriot with a man of science. Thus, collecting stones, drawing monuments, picking herbs on the mountain or turning archives upside down is seen by the common people as being synonymous with being a Catalanist' (Prat de la Riba 1906a).

Alcover himself, in addition to his opening and closing speeches, read a short paper: *Concordansa del participi ab el terme d'acció* ('Participle agreement with the action term') (1908a) and a long report: *La llengua catalana té sintaxis pròpia* ('Catalan has its own syntax', 1908b). Among the numerous foreign linguists in attendance, the previously mentioned Saroïhandy stood out with a communication on Pyrenean Catalan in contact with Aragonese (1908). Of the domestic attendance, Pompeu Fabra read an outstanding paper: *Qüestions d'ortografia catalana* ('Questions of Catalan spelling', 1908) and took active part in numerous debates in which he put forward amendments with impressive arguments when the subject required it, as was the case for Alcover's papers (Julià-Muné 2006).

3.1.2 Alcover, as a successful dialectologist and a frustrated grammarian

Alcover's works encompass both spelling and pronunciation and constitute the basis for the birth of Catalan dialectology by means of 'philological fieldwork' ('eixides filològiques' in 1906 and in 1921) (Perea 2005; Moll 1983b). One of Alcover's constant preoccupations was the study of the pronunciation of the different varieties of Catalan and of Latin (Alcover 2004), which he considered an essential task in linguistic studies. With this in mind, he expressed his opinion in August 1902: '[Philology] was born largely from the deep and detailed study of popular pronunciation. [...] and the formulation, explanation and demonstration of these laws [of pronunciation] constitute the main body of linguistics' (Alcover 1902b, 141–142).

Another aspect that particularly interested the Majorcan philologist was necessarily derived from the previous one: the sound-grapheme relation. He shared this interest with his political mentor Prat de la Riba, who responded with his spelling-related concerns in a letter to Alcover (10th Oct., 1904): 'You understand that although in the dictionary all spellings are registered, the one adopted by the dictionary itself is of great importance and if all the means of diffusion of Catalan, the main daily newspapers, schools, the Bible, the dictionary, adopt the same spelling, we will have accomplished our main task of fixing it definitively' (Julià i Muné 2000b, 61–62). Alcover replied by reminding him of the fundamental role that Prat's newspaper could play in disseminating standardised spellings (14th Dec., 1904).

Alcover has been widely discussed as a dialectologist and lexicographer and is known – and recognised – well enough from his more characteristic works in dialectology (Alcover 1909; Moll/Alcover 1929–1932; Perea 1999a; 1999b) and lexicography (*DCVB*). His extensive review from 1909, in which he summarises and criticises in detail Morel-Fatio's (1888) and Saroïhandy's (1904) contributions to Catalan phonetics in the previously mentioned chapter *Das Catalanische* in the first and second edition, respectively, of Gröber's *Grundriss der romanischen Philologie*, consecrates him as the founder of Catalan dialectology. The other two works, as we will see later, are projects that he could not print personally. The first, on *verbal flexion*, was ultimately completed and edited by Francesc de B. Moll and Pilar Perea, while the second, the project of an entire lifetime, was a dictionary, that was later completed by Moll (see 5) with the collaboration of Manuel Sanchis Guarner (1911–1981) and his daughter Aina Moll.

Alcover's dialectological research was based on frequent fieldwork to collect data, known as *eixides dialectològiques* ('dialectological excursions'). We know from his own records that the longest of these excursions were made during the summers of 1906 and 1921. For the first one he collaborated with Bernhard Schädel (Moll 1983b), whom he met at the Perpinyà (Perpignan) railway station on 3rd August 1906, according to Alcover's account of the *eixida* in his journal: 'Dr. Schädel, getting up so early, so fit, in high spirits, so enthusiastic as ever for our philology. We hadn't seen each other for two years; we used to keep in touch by letter. [...] He tells me what he

thinks about that research and shows me a list of about seventy irregular verbs in a book he shows me [...]’ (Alcover 1907, 260).

On the second *eixida*, fifteen years later, Alcover was accompanied by a promising young man of only seventeen years of age: Francesc de B. Moll, who gained his first experience as a field researcher in dialectology under Alcover’s guidance (Moll 1970, 134–140). According to Alcover’s account (1922, 225), that summer of 1921 young Moll met the industrialist and linguist Alfons Par (1879–1936) from Barcelona, the future author of *Sintaxi catalana segons los escrits de Bernat Metge (1398) (1923)* (see 3.1.4) in the village of Sort. Moll would in due course become Alcover’s heir and great collaborator.

Was Alcover a frustrated grammarian? As far as grammatical studies are concerned, Alcover’s contribution to grammar (phonetics, morphology and syntax) constitutes an incomplete project. He undoubtedly wished to complete Forteza’s unfinished grammar by adding the missing syntax (see 2), but in the end he merely edited that grammar respecting the author’s original work (Forteza 1915). However, Alcover’s grammatical contributions were considerable (Julià-Muné 2005a). In fact, he introduced grammatical studies with *Questions de llengua y literatura catalana* (1903), a long article of 350 pages in response to Ramón Menéndez Pidal’s *Cataluña bilingüe* (1902), in which the Spanish linguist stated that Catalan was little more than a variety of Spanish, without its own grammar. Preparing this extensive response represented the first stage of Alcover’s autodidactic linguistic education, after the introduction he had received from his teacher Tomàs Forteza some years before.

3.1.3 Pompeu Fabra’s linguistic works and his descriptive grammar for L1 and L2 learners (1912)

As Robins informs us, “The principal and most obvious contrast between the last two centuries has been the rapid rise of descriptive linguistics, as opposed to historical linguistics, to its present position of predominance. This has become the source from which the major developments in contemporary linguistics have sprung; [...]” (1967, 199). Therefore, Fabra’s grammatical and linguistic (or proper scientific) studies were produced between two great periods in the history of language science, those of Neogrammarian and structuralist linguistics.

In the second decade of the 20th century, and afterwards, Fabra’s priority was the institutional codification of the language and most of his linguistic production follows Meyer-Lübke’s much-admired Neogrammarian approach as well as the first works written by Saussure, Jespersen and the Prague School. His linguistic work³ is actually included in his descriptive grammars, especially in the one written outside Catalonia

³ More detailed information may be found in Mascaró (2006).

and published in 1912, which was the best suited for L1 and L2 learners of the language at the time. This influence can also clearly be seen in a series of articles that would pave the way for the codification of Catalan (Fabra 1897; 1903; 1905; 1906; 1907; 1908; 1913; 1914; 1926).

3.1.4 Arteaga and the IPA, Par and the syntax and Calveras and the norm

Josep M. Arteaga Pereira (1846–1913) was a composer from Barcelona who made a name for himself with his paper at the First Congress: *Ullada general a la fonètica catalana. El seu caràcter propi dins la família novo-llatina* (1908). He was a member of the International Phonetic Association and contributed regularly (1904) to its magazine *Le Maître Phonétique*. In fact he had become such a prominent phonetics specialist at the time up to the point that Fabra asked him to revise the first chapter of his 1912 grammar. Pere Barnils edited Arteaga's works that were published by the IEC (Arteaga 1915).

Other relevant contributions⁴ include those of Alfons Par (1879–1936) as an active collaborator within the 'Alcoverian galaxy', who represented the IEC in Great Britain in 1909 but produced his work on syntax outside of any institutional framework: *Sintaxi catalana segons los escrits en prosa de Bernat Metge (1398)* (1923). Josep Calveras (1890–1964) in his *La reconstrucció del llenguatge literari català* (1925) was especially critical of the officialdom in Catalan studies and advocated for a more prominent role for spoken Catalan as an indispensable foundation for developing a proper standardised language.

3.2 Institutionalised linguistics (1913–1925)

In the early 20th century three institutions were founded in an almost Russian doll-like structure: the Institut d'Estudis Catalans (1907), the Secció Filològica (1911) and the Oficines Lexicogràfiques (1912). These institutions would prove essential to advancing linguistic research in Catalonia.

3.2.1 The German influence on Catalan linguistics and the first project for studying linguistics abroad: Romance linguistics at the Prussian University of Halle

The young postgraduate Bernhard Schädel (1878–1926) met Antoni M. Alcover in Majorca in 1904. The Balearic linguist became Schädel's first teacher of Catalan and,

⁴ For a full account of Par's and Calveras's contributions, see Ribes (2011) and Iglésias (2004) respectively.

together in close collaboration, they undertook a great deal of fruitful research in Catalan linguistics, including a number of ambitious projects: a) an international congress on syntax, which ended up focussing on linguistics as a whole (the First International Congress of Catalan); b) a new linguistics journal specifically on Catalan; c) a new grammar; d) the dictionary; e) a research institute, and f) promoting and funding scholarships for young Catalans to study Romance linguistics at Halle University. Four of the projects were quite successful: a, b, e and f, that is, the congress, the new journal (see 3.2.2), the IEC and the three scholarships funded by the Diputació de Barcelona. As part of the latter, four years after Schädel's first visit to the Balearics, three young Catalans won a scholarship to study in Halle, where Schädel had been lecturing in Romance languages, including Catalan. These three scholars were also the first native speakers to work as assistants (later known as *lectors*) helping to teach Catalan abroad. Their research and teaching work was supervised by Schädel over a two-year period from 1908 to 1910 (Julià-Muné 2008). The scholarship holders, known as *pensionats*, from the old-fashioned term *pensió*, or *estipendiats*, from the German *Stipendium*, at the time, were Manuel de Montoliu (1877–1961), Pere Barnils (1882–1933) and Antoni Griera (1887–1973).⁵ This starting point was the basis for the wide international expansion of Catalan studies outside the Catalan-speaking regions up to the late 1930s and especially from the mid-20th century onwards in post-war Europe and North America.⁶ Schädel did not write a Catalan grammar but did produce the first, if brief, *Manual de fonètica catalana* (1908), aimed primarily at helping to transcribe the linguistic materials collected by Alcover's collaborators for his dictionary in progress.

3.2.2 The Lexicographical Offices of the IEC and its sections

Fabra became director of the Oficines Lexicogràfiques of the IEC in Barcelona as soon as they were founded on 11th October 1912, just after he was recruited by President Prat de la Riba. According to the *Reports* published by the Offices (Fabra/Barnils 1915; 1917) and the *Guia de les Institucions de la Diputació de Barcelona* (1916, 13–33), the Offices were in fact two separate departments: the Oficina del Diccionari General de la Llengua Catalana, run by Pere Barnils, delegated by Alcover – at the time Head of the Secció Filològica –, and the Oficina de l'Inventari Aguiló, run by Fabra himself, who was also General Director of the Offices.

⁵ A detailed account of the German-Catalan relationship over more than a century can be found in Pons/Skrabec (2008). Schädel's influence on Catalan linguistics can be traced in Moll (1965) and Julià i Muné (2000b).

⁶ See Bover (1993) for further information.

The main goal of the Oficina del Diccionari General was the creation of the dictionary initiated by Alcover. Between 1915 and 1918 the Office had all of Alcover's archives, known as *la Calaixera* ('the chest of drawers'), which had been transported from Palma to Barcelona, at its disposal. In order to coordinate all the activities of the Office, particularly with regard to Alcover's collaborators, a new journal was founded: the *Butlletí de Dialectologia Catalana* (1913–1937).⁷

The Oficina de l'Inventari Aguiló, was in charge of ordering, classifying and editing Marià Aguiló's legacy, supervised by his devoted disciple Pompeu Fabra, assisted by Manuel de Montoliu. The Office was responsible for the partial publication of the *Diccionari Aguiló* (1915–1934). For a period of time (1918–1924) work was also done on the official dictionary of the IEC itself, known as *Diccionari de la llengua literària*, but it was never published (Rafel i Fontanals 1996b). In due course the Office also edited the *Diccionario Balari* (1926–1936). Unfortunately, about 50 % of Josep Balari's materials remained unpublished due to the outbreak of the Spanish Civil War (see 2).

3.2.3 The birth of experimental phonetics in Spain (1913–1921): Pere Barnils and his *Estudis fonètics* (1917)

Pere Barnils introduced experimental phonetics to Spain by founding the country's first laboratory for developing this new branch of linguistics together with an accompanying journal on phonetics. Barnils established the foundations of Catalan phonetics as a science and he was the first to apply experimental techniques in order to correct speech and voice disorders.⁸ He was also a regular contributor to international conferences and journals (Barnils 1912).

In 1913, after returning to Barcelona from Halle and Paris, where he had been introduced into dialectology by Schädel and Suchier and into experimental phonetics by the Abbé Rousselot, he edited the first journal of Catalan linguistics, the *Butlletí de Dialectologia Catalana*, which turned out to be the first periodical on modern linguistics to be published on the Iberian peninsula.⁹ At the same time he started to organise what was to become the first Laboratory of Experimental Phonetics in Spain (1913–1921).¹⁰ It

⁷ More detailed information on the origin and initial development of this journal can be found in Julià i Muné (2000c: 125–130).

⁸ For an overview of experimental phonetics studies in Spain, as developed in Barcelona and Madrid, see Julià-Muné (2010).

⁹ After Pere Barnils's tenure, the editorship of the *Butlletí de Dialectologia Catalana* (1913–1937) would be successively assumed, in order, by Antoni Griera, Pompeu Fabra and Joan Coromines.

¹⁰ The following year, at the Centro de Estudios Históricos, in Madrid, Tomás Navarro Tomás founded a phonetics laboratory, which was active for a much longer period and was only cut short by the breakout of the Spanish Civil War.

was set up at the Lexicographical Offices of the IEC in the Palau de la Diputació in Barcelona, the present-day headquarters of the Generalitat de Catalunya, and it was officially inaugurated in 1914. By 1915, the laboratory was working on a number of significant projects and was running full-time from mid-1916 up until the end of 1917.

In April 1914 Barnils represented Catalonia at the First International Congress of Experimental Phonetics held in Hamburg and was elected as the representative for Spain in the newly created International Association of Phonetics. For a few weeks he worked with Panconcelli-Calzia at his laboratory housed at the Hamburg Colonial Institute.

During World War I, which broke out the following summer, Barnils's laboratory became, almost by accident, the leading centre in experimental phonetics in Western and Southern Europe. Under Barnils's leadership, the laboratory became a major scientific research centre with an interdisciplinary approach; in addition to conducting experimental studies of Catalan descriptive phonetics, he worked with speech therapists and welcomed all sorts of researchers – from both the Catalan-speaking regions and from abroad – whose interests were related in some way or other to phonetics. In the summer of 1917 he published his magnificent first – and last – volume of *Estudis fonètics*, which included projects carried out at the laboratory, such as studies on the articulations of /k/ and /g/ in Majorcan Catalan, vowel nasality, alveolar roll vibrations and articulatory force in voiceless plosives (to cite only those conducted by Barnils in Catalan phonetics). The following year Barnils submitted a detailed report (1918) on the extensive research activities undertaken at the laboratory. Unfortunately he had to leave his experimental work both for political reasons and due to conflicts with management. As a result, his first phonetics laboratory was shut down in 1921, after three years of inactivity (Julià i Muné 1984; 2000c).

Under Badia i Margarit's guiding, Ramon Cerdà (1972) tried to follow the path set by Barnils, now at Barcelona University, and the IEC resumed its work at the end of the century under Daniel Recasens's updated research work (see 8.1) with fruitful results (1986; ²1996).

3.2.4 Fabra as a language reformer: the orthoepy/orthography and de facto prescriptive grammar and dictionary

'You have to bear in mind that at the start of this [literary] Renaissance our writers had at their disposal an impoverished, distorted and spoiled language that contained a large amount of utterly unnecessary loanwords from Spanish. They could not adopt such a language as a literary language before purifying and enriching it in order to elevate its status. And this could not be carried out without a perfect knowledge of the old language and the present dialects, which might guide us through the most difficult task of identifying and subsequently correcting the deviations suffered by our language' (Fabra 1932a, 22)

After collaborating with the Grup de l'Avenç for three decades and participating in the First International Congress of Catalan, in 1915 Fabra focused his presidential speech, in that year's *Jocs Florals* ('Poetry Contest') in Lleida, on the defence and promotion of the minoritised Catalan language, as quoted above. Fabra insisted upon the urgent need to enhance Catalan as a literary language by cleansing and improving it, so that once having recovered its genuineness, it could become what it is known as a 'standard language' fit for purpose in the present day.¹¹ Not much later, owing to the favourable political climate prevailing at the time, he would gradually publish his main works, which, albeit unintentionally, were to become the prescriptive corpus of the Catalan language, that is, the three-fold codification foundation – orthographic, grammatical and lexicographical – of Catalan: *Diccionari ortogràfic* (1917), *Gramàtica catalana* (1918) and *Diccionari general de la llengua catalana* (1932b), respectively.¹²

4 An intermittent balance (1926–1939)

At the end of the first quarter of the century a great feat was achieved: Wilhelm Meyer-Lübke consecrated, so to speak, Catalan as a Romance language in *Das Katalanische*, published in 1925. This work was translated into Catalan by the end of the century (Calaforra 1998).

However, the balance did not remain steady. The dictatorship of General Miguel Primo de Rivera (1923–1930) brought out the dissolution of the Mancomunitat de Catalunya in 1925 and put the IEC out of operation due to the repression of Catalan language and culture. This precarious situation continued until the Second Spanish Republic was proclaimed in 1931 and the Generalitat de Catalunya (Autonomous Government of Catalonia) was restored and the Estatut d'Autonomia was approved in 1932. However, the short-lived October revolution of 1934 sent Catalonia back to its pre-Republican state until February 1936 and the return of a Spanish government favourable to Catalonia's interests, five months before the outbreak of the Spanish Civil War. The Universitat Autònoma de Barcelona, founded in 1933 and where Fabra lectured on historical Catalan grammar and advanced Catalan language, was closed down between November 1934 and February 1936, after being open for a mere three semesters.

¹¹ For a detailed account of this process, see Lamuela/Murgades (1984).

¹² Fabra's work devoted to the corpus planning or codification of Catalan is included in his *Obres Completes* ('Collected Works'), edited by Jordi Mir and Joan Solà (Fabra 2005–2013, 9 vol.).

4.1 The *Diccionari català-valencià-balear* and verbal inflexion materials begin to be published

After the dramatic reduction in the activities of the IEC, linguistic work was partially taken over by research centres such as the Reial Acadèmia de Bones Lletres de Barcelona – which published one of Badia i Margarit's early works (1950) – or the Biblioteca Balmes via its official organ, the *Anuari de l'Oficina Romànica de Lingüística i Literatura* (AORLL), which included contributions such as Moll (1928–1931) and Moll/Alcover (1929–1932). Even Spanish projects such as Navarro Tomás's *Atlas Lingüístico de la Península Ibérica* (ALPI, 1962) involved the collaboration of linguists specialised in Catalan, such as the Minorcan Francesc de B. Moll (1903–1991) and the Valencian Manuel Sanchis Guarner (1911–1981), with fruitful results (Navarro Tomás/Sanchis Guarner 1934). This period also saw the publication of the results of the most relevant research project in descriptive lexicography to be undertaken for a quarter of a century, the *Diccionari català-valencià-balear*. In 1926 Alcover and young Moll published the first parts of their dictionary in Palma de Mallorca and in 1930 the first volume appeared. From then on Moll took over many of Alcover's tasks, including editing the *Bolletí del Diccionari de la Llengua Catalana*, the *Diccionari* – the last volume of which appeared in 1962 – and the majority of an ambitious dialectological and morphological project, *La flexió verbal en els dialectes catalans* (Moll/Alcover 1929–1932), which had been carried out by Alcover. The fully edited version of the latter was published fairly recently (Perea 1999a; 1999b).

4.2 Language and applied linguistics

Pere Barnils's interest in vocal and speech disorders led to the founding of his private *Laboratori de la Paraula* (1914–1933) – with its journal *El Parlar* (1931–1932) – and to Barnils taking charge of the Barcelona School of Deaf-mutes (1918–1930) to which he devoted almost the entire rest of his life. There he founded his third laboratory of experimental phonetics and edited its journal *La Paraula* (1918–1921). As a result, his most significant publication in the field of logopaedics was *Defectes del parlar* (1930).

Among applied linguists devoted to pedagogy, Alexandre Galí (1886–1969) must be mentioned due to the significance of his works: *L'ensenyament de l'ortografia als infants* (1926), *Lliçons de llenguatge* (1931a), *Per la llengua i per l'escola* (1931b) and *Introducció a la gramàtica* (1935).

4.3 The *Converses filològiques* (1919–1928) and Pompeu Fabra's lexical reform (*Diccionari general de la llengua catalana*, 1932)

In spite of the unstable political situation, Fabra continued to publish his 'columns about language' in the periodical *La Publicitat* from 1919 until 1928. Today they are collected in Volume 7 of his *Obres completes* (2010). Moreover, after the *Diccionari de l'Institut* was temporarily set aside and the Oficines were shut down in 1924, Fabra started publishing his provisional dictionary, which would later become the prescriptive *Diccionari Fabra*, in instalments. As he states in the preface of its first edition: 'The work we were entrusted with – which appears now – can actually be considered as an outline of the future Dictionary of the Institute. [...] a dictionary as the present one which aims at being prescriptive [...]' (Fabra 1932b, VI).

4.4 Moll and Joan Coromines: journals forced to close

In 1937 the two journals regularly published in Catalan had to close down definitely. Moll brought the second period of Alcover's *Bolletí del Diccionari de la Llengua Catalana* to an end in Palma de Mallorca, while in Barcelona Joan Coromines (1905–1997) did his best to edit the final 24th volume of the *Butlletí de Dialectologia Catalana*. Fortunately, some of his incipient works on historical morphology and dialectology during the last stage of this period of Catalan linguistics did manage to see the light: *El parlar de Cardós i Vall Ferrera* (1936) and *Mots catalans d'origen aràbic* (1937).

5 Post-war and exile linguistics (1939/1940–1967)

After Franco's victory in the Spanish Civil War, the struggling field of Catalan linguistics was resumed in exile by Fabra in France (1939–1948) – with new grammars published in Paris (1941; 1946) and his posthumous one, edited by Coromines, in Barcelona (1956) – and Coromines himself (1939–1967), who began to produce a prolific body of work in Argentina and in the US, specifically at the University of Chicago, while researching and lecturing in Romance linguistics, in historical grammar and in Spanish and Catalan lexicography. Coromines edited the long-awaited *Miscel·lània Fabra* in Buenos Aires (1943), while in Britain Joan Gili published his *Introductory Catalan Grammar* (1943). In 1967 Coromines came back to Catalonia and this period of Catalan linguistics was over.

In the Catalan-speaking territories new researchers, Antoni M. Badia i Margarit (1920–2014) in Barcelona (1951) and Francesc de B. Moll in Madrid (1952), published the indispensable Catalan historical grammar based on Ramón Menéndez Pidal's manual for Spanish (1904). A reading of their works would not be complete without Coromines's historic and extensive review (1958). Later on a new Catalan grammar

describing “el catalán literario moderno” (Badia i Margarit, 1962) was published in Madrid. During this period, a quite different linguistic development took place, namely the 7th International Congress of Romance Linguistics, held at the University of Barcelona in 1953 under the supervision of Antoni Griera, the official linguist of the prevailing political regime. In addition to Moll, Coromines and Sanchis Guarner, Badia i Margarit and Germà Colón stood out among the youngest contributors. Two additional works of Fabra’s were published posthumously: the second edition of the *Diccionari General* (1954) and a selection of the *Converses filològiques* in ten brief volumes (1954–1956). A few years later Badia diffused an influential essay on Catalan language and culture (1964) (see 7.10).

The resumption of the linguistic studies in the Valencian Country and in the Balearic Islands was especially significant with Sanchis Guarner publishing his grammar in Valencia (1950) and assisting Moll in finishing the *Diccionari català-valencià-balear* (DCVB, 1926 in parts; 1930–1962) in Palma. In 1968 the latter appeared in standard spelling.

6 The dawn of modernity (1968–1977)

In the late 1960s and early 1970s Catalan linguistics, after having obviously lost momentum over the preceding three decades, was reinvigorated with a new enthusiastic impulse and was about to enter its most prolific period. Fabra’s centenary (Aramon 1963–1968) and the Strasbourg conference on Catalan linguistics were the forerunners of a promising new era, the *Associació Internacional de Llengua i Literatura Catalanes* was born in this period and Francesc de B. Moll brought Alcover’s great lexicographical project to completion.

6.1 Fabra’s centenary and the Strasbourg conference: the *Associació Internacional de Llengua i Literatura Catalanes* (AILLC)

It should be noted that the first post-war international conference on Catalan was held in France (*La linguistique catalane*, Strasbourg, 1968), the second at the University of Amsterdam in 1970 and the third in 1973 at Fitzwilliam College, Cambridge, where the AILLC was set up.¹³

¹³ Henceforth an international conference funded by the AILLC has been held every three years alternating between host institutions in the Catalan-speaking regions and abroad. The University of Bucharest hosted the most recent conference in 2018.

The proceedings of that first post-war Catalan conference were published in French in Paris (Badia i Margarit/Straka 1973) five years later. They included five papers written by Ramon Aramon i Serra (1973) on the history of the Catalan language, Joan Solà (1973) on orthography and grammar, Germà Colón (1973) on vocabulary, Joan Veny (1973) on dialectology and Henri Guiter (1973) on onomastics, as well as three more papers produced by the Catalan co-organizer Antoni M. Badia i Margarit (1973a; 1973b; 1973c) on descriptive phonetics and phonology, morphosyntax, and the sociolinguistic position of Catalan at the time, respectively.

6.2 Linguistics in the Balearic Islands and the Valencian Country

Moll published a grammar (²1968; ¹1937) and promoted the Obra Cultural Balear, which contributed to him holding the first chair in Catalan at the extension of the University of Barcelona located in the Balearics in the early 1970s. During this period Sanchis Guarner updated *La llengua dels valencians* (⁴1972, ¹1933) in Valencia and paralleled Moll's enthusiasm within a Valencian triad of cultural promoters, Joan Fuster, Manuel Sanchis Guarner and Enric Valor.

6.3 Criticism in Romance and general linguistics and the revival of syntax

Another Valencian, from Castelló, Germà Colón (1928–) proved himself to be an acute critic in the field of Romance linguistics while teaching at the University of Basel. In 1976 he published *El lèxic catalán en la Romania* and organised the Fourth Meeting of Catalan Language and Literature (see 6.1) in the Swiss city. By matching Colón's criticism, especially in general linguistics, and lecturing heterodoxically at the recently inaugurated Universitat Autònoma de Barcelona (1968), Gabriel Ferrater (1922–1972) developed a strong following in his short lifetime (Murgades 1972; Argenter 1972; Mascaró 1984). His works, collected in the volume *Sobre el llenguatge* (1981), though brief, are sharply critical and incisive. A third young university lecturer, Joan Solà (1940–2010) began publishing his first works on syntax (1972–1973)¹⁴ as well as on language correctness, the latter dealing with crucial works produced over the last five centuries (1977a).

¹⁴ This two-volume publication was expected to be prefaced by his close friend and colleague at the reborn Universitat Autònoma de Barcelona, Gabriel Ferrater. The preface was never written due to Ferrater's death a few months before the publication of Solà's work.

6.4 Catalan linguistics beyond the Pyrenees

Northern Catalan, spoken in Southern France, was the specialism of Pere Verdaguer (1929–2017), who emigrated there from Northeastern Catalonia in 1939. He studied this variety of Catalan (1974a, 1974b) and conducted research on contrastive linguistics with a specific focus on French-Catalan interferences (1976). He was a dynamic promoter of Catalan culture and founded the Grup Rossellonès d'Estudis Catalans and the Universitat Catalana d'Estiu in Prada de Conflent.

6.5 Joan Coromines: revisiting linguistics essays

Back in Catalonia, Coromines contributed to Fabra's Centenary celebrations in Barcelona and Paris (1968), and he published the second volume of his toponymic studies (Coromines 1965–1970) and the volume *Lleures i converses d'un filòleg*, which was awarded the Lletra d'Or prize in 1971. It included, among others (1958; 1971a), his pioneering article on Catalan orthology¹⁵ *Sobre l'elocució catalana*:

'Dramatic elocution must be a little bit more distinguished than the familiar type. However, it should not be an affected pronunciation. At least some considerations must be taken into account to differentiate it from Barcelona pronunciation: you must always fully voice the combination *tj* (or *tg*) –*fetge, sitja, jutjar*... –, as well as *bl i gl* in words such as *poble, reblar, estable, regla, joglar*. Do not leave out *e* (or *a*) next to *r*, in words such as *però, veritat, feredat, escarabat*; nor final *-t* or *-d* after *r* in *curt, tort, verd*; do pronounce *vull* and *cella*, but not *vui* nor *ceia*; etc.' (Coromines 1963: 340).

Later on, he published the three-volume *Entre dos llenguatges* (1976–1977), a collection of his most relevant works on historical phonetics.

May 1974 saw the publication of a new linguistic and literary journal, *Els Marges*,¹⁶ which was founded by the dynamic academic and literary critic Joaquim Molas. It should be noted that the *Butlletí de Dialectologia Catalana* had closed down in 1937 and been to a certain extent replaced by the *Estudis Romànics* and *Serra d'Or* in the late 1940s and mid-1950s, respectively. However, *Els Marges* was a new magazine for a new era. It played a crucial role in recovering the spirit and goals of the old *Avenç* with its innovative as well as combative stance favouring Catalan culture. Later on, a new linguistics journal, *Caplletra*, was launched in Valencia. It seems appropriate, then, to close this period of Catalan linguistics with Badia i Margarit's remarks about precisely the peculiarities of Catalan culture:

¹⁵ A more detailed account of Fabra's and Coromines's contribution to Catalan orthology may be found in Julià-Muné (2012).

¹⁶ For a detailed description of its first 25 years, see Viana (1999).

‘Catalan culture can be compared with any other culture, if we take into account the value of its contributions or the topics it deals with or the methods it applies, and so on. From that point of view our culture can be objectively evaluated. Nevertheless, considering the way it usually develops, it is noticeable that Catalan culture lives in permanent tension because of conditioning elements. As a result, our culture is not comparable with other models any more. Catalan culture displays a twofold sign of identity that makes it extremely peculiar, even unique. Science cannot be kept apart from passion’ (Badia i Margarit 1977, 5).

7 Contemporary linguistics I: the revival of home rule (1978–2000)

After Franco’s death, Catalonia and the other Catalan-speaking territories in Spain enjoyed a certain amount of freedom and democratic home rule. However, despite the restoration of democracy, two basic pillars of the Francoist status quo remain intact, the monarchy, after the dictator appointed Juan Carlos de Borbón as his successor, and the ‘national’ unity of Spain (see Appendix). This period has, however, proved the most fruitful as far as research production in linguistics is concerned and extends well beyond the turn of the new millennium, both in the Catalan-speaking territories and abroad, especially in German- and English-speaking countries.¹⁷ These opposing trends in the position of Catalan are reflected in two key events that both took place in 1978: the establishment of the Spanish Constitution and the promotion and founding of the Institut de Filologia Valenciana by Sanchis-Guarner. In 1994, this became the Institut Interuniversitari de Filologia Valenciana (IIFV).

7.1 Historical lexicography, historical linguistics and onomastics

Historical lexicography and onomastics in this period were clearly led by Coromines, who had been gathering data for decades, as shown in his main dictionaries, the *Diccionari etimològic i complementari de la llengua catalana* (1980–2001) and the *Onomasticon Cataloniae* (1989–1997). At the University of Basel, Germà Colón worked on Catalan lexical units of Romance origin (1976), based on different kinds of historical texts (1978), as well as contrasting Catalan and Spanish (1989). In Catalonia and Valencia several lexicography historians were at work, including Bruguera (1985), Colón and Soberanas (1991; 1986), Rico and Solà (1995) while others approached the

¹⁷ A sample of the expansion of Catalan studies abroad may be found in the works of linguists such as Gili (1943) and Roca-Pons (1971) in grammatical studies; Gulsoy (1982; 1993) and Rasico (1982; 1993; 2006) in historical grammar; Russell-Gebbett in Old Catalan texts (1965), and Wheeler in phonology (1979). In Spain, Alarcos Llorach’s contribution (1983) stands out in particular.

problems of Catalan historical grammar from a generative point of view (Duarte/Alsina 1984). Joan Martí i Castell discussed problems and methods of historical grammar (1990) and Coromines's disciple par excellence, the Canadian of Turkish origin Joseph Gulsoy, published some results of his research on this topic (1993), as did a Catalan disciple, Moran (1994). An American scholar followed suit working on the phonology of Old Catalan (Rasico 1982; 1993) and Badia i Margarit (1999) had resumed his interest in the medieval text *Regles de esquivar vocables*, as did Colón at the turn of the century (2001).

Meanwhile the Lexicographical Offices of the IEC had issued their new prescriptive dictionary (1995), which was updated twelve years later (2007).

In onomastics, Enric Moreu-Rey (1917–1992), the founder of the Societat d'Onomàstica issued his work on anthroponymy (1993) and Josep M. Albaigès published his popular dictionary of personal names (¹2000, ¹1980). Moll released his new version of “llinatges catalans” (²1982; ¹1959), and Josep Moran (1995) and Joan Miralles (1996) directed their focus onto onomastics in their historical studies.

7.2 Geolinguistic variation: the Catalan dialects

The results of linguistic variation studies (Lloret et al. 1997) established geolinguistic variation or traditional dialectology as the most cultivated branch of traditional Romance linguistics. Catalan dialectology was fathered by Alcover, but we can attribute the current high standards of the discipline to another Majorcan, Joan Veny (1932–). Veny ushered in the last period of 20th-century Catalan linguistics with his successful *Els parlars catalans* (¹²1998; ¹1978a), which he followed up with geolinguistic studies (1978b), an introductory manual (1986) and an extensive sampling of geolinguistic texts, co-authored with Lúdia Pons (1998), which announced the *Atles lingüístic del domini català* (Veny/Pons 2001–2017) published by the IEC.

Moll, as Alcover's heir and direct collaborator had a thorough knowledge of the Majorcan dialect (1980) as did Colomina (1985; 1999) regarding the Valencian Community and Aragonès (1995) concerning Southern Catalonia, together with a handful of Veny's active disciples (e.g. Corbera, Pradilla, Bibiloni, Sistac).

7.3 Phonetics, phonology and morphology

The leading phonetician of the period was Daniel Recasens whose works encompass experimental phonetics (1986), descriptive and geolectal phonetics (²1996; ¹1991) as well as the manual *Fonètica i fonologia* (1993). Badia i Margarit collected his own phonetic studies in *Sons i fonemes de la llengua catalana* (1988) and Miquel Àngel Pradilla edited *El món dels sons* (1998), while Eulàlia Bonet, Maria-Rosa Lloret and Joan Mascaró issued a manual on phonetic transcription (¹1997; ²2000). Two pronun-

ciation dictionaries were published: Jordi Bruguera's *Diccionari ortogràfic i de pronúncia* (1990, updated in 2004 to include some onomastic terms) and the *Diccionari de pronunciació del català* by David Paloma and Albert Rico (2000). All of them dealt with segmental aspects of Catalan. Suprasegmental features, especially those regarding intonation and metrics have been studied by Pilar Prieto (Prieto/Cabré 2013) and Salvador Oliva (1992), respectively (see 8.3).

In phonology, relevant works were published by Palmada (1994) and by Bonet/Lloret (1997), the latter author acting as co-editor of the phonetics and phonology chapters of the Catalan descriptive grammar (*GCC*) that was published in the new century (2002). In phonology and morphology, the works of Mascaró (1978; 1986), the co-editor of the chapters on morphology in the *GCC*, hold a prominent position alongside Wheeler's *Phonology of catalan* (1979).

7.4 Syntax

A manual on generative syntax (Bonet/Solà 1986) was published, and soon afterwards Joan Solà, while working on the prolific contributions of Fabra (Solà 1987a) and Coromines (Solà 1999), presented the results of his studies on both descriptive (1987b) and prescriptive (1994) syntax. Later on Solà promoted an ambitious project (see 8) for a new Catalan descriptive grammar (Solà et al. ¹2002; ⁴2008), which was co-edited by Lloret, Mascaró and Manuel Pérez Saldanya, the latter also working on morphosyntax (1988; 1998). M. Josep Cuenca, following Solà's path, contributed considerably to the spread of Catalan syntax (1988–1991).

7.5 Grammar and standardisation

Gemma Rigau, working on generative grammar, was one of the most regular contributors in the first period of *Els Marges* (1975). Later on, she published a discourse grammar (1981) and edited the results of work carried out in collaboration with Joan Mascaró, Anna Bartra and Josep M. Nadal (1984). A dictionary of verbal usage (Ginebra/Montserrat ¹1999, ²2009) included information on structures linking every verb and its variants to their complements. Three grammars also appeared in the late 20th-century: Hualde published his in an updated linguistic framework (1992), Badia i Margarit tried to design an ambitious descriptive and prescriptive grammar (1994), and Wheeler, Yates and Dols (1999) published a comprehensive work that aimed to provide an up-to-date systematic description of modern standard Catalan by acknowledging regional varieties.

In 1983 the University of Barcelona promoted conferences to evaluate the standardisation of the language (Cabré et al. 1984; 1987; Martí/Pons/Solà 1989) and Francesc Vallverdú (1935–2014), following in Coromines's footsteps, was the first to propose

criteria for elocution and orthology in Catalan (1986). During the following decade the IEC proposed standardising spoken Catalan (1990b; 1992), under the guidance of the president of its Philological Section, Antoni M. Badia i Margarit, while Solà (1990) and later on Rossich and Rafanell (1998), among others, contributed to discussion of this as yet unsolved question.

7.6 Semantics, lexicology, terminology and computational linguistics

Semantics and lexicology have been worked on extensively by M. Teresa Cabré and Gemma Rigau (Cabré/Rigau 1985). The first author was head of the Centre de Terminologia Catalana (TERMCAT) and later of the Institut Universitari de Lingüística Aplicada ('Applied Linguistics University Research Institute' (IULA)) at Pompeu Fabra University, where she carried out her research work on the theory, methods and applications of terminology (1992). Later on, Cabré promoted and coordinated research projects on neology (see 8.3).

After having devoted himself to dialectology, phonetics and lexicography, Joaquim Rafel i Fontanals edited the *Diccionari de freqüències* (1996a; 1998a; 1998b) and has been coordinating an ambitious IEC project, the *Corpus Textual Informatitzat de la Llengua Catalana* (see 8.1).

7.7 National history and language

The Valencian Sanchis Guarner, the author of the *Aproximació a la història de la llengua catalana* (1980), left his project on the history of the Catalan language uncompleted. Two years later Josep M. Nadal and Modest Prats issued the first volume (1982) of the history of Catalan, covering the development of the language from its early origins up until the beginning of the 15th century. The second volume covering the 15th century appeared in 1996. In the meantime, Solà published his studies on language history (1991) while Nadal released his incisive *Llengua escrita i llengua nacional* (1992). In 1995 Balsalobre and Gratacós edited *La llengua catalana al segle XVIII*.

At the turn of the century, the Balearic Joan Miralles published his studies on language history (2001). However, it was in Valencia that a synthetic history of Catalan language was published at the beginning of the new century (Ferrando/Nicolàs 2005), a quarter of a century after Sanchis Guarner published the first part of his history.

7.8 Functional variation, discourse analysis and pragmatics: colloquial Catalan

Lluís Payrató at the University of Barcelona, Margarida Bassols at the Universitat Autònoma de Barcelona, Josep M. Castellà at Pompeu Fabra University and Josep Lacreu at the Valencian Academy of Language are the leading linguists in the field of discourse analysis and pragmatics in Catalan. Payrató issued the innovative *Català col·loquial* in 1988, and also edited and co-edited two key collective works, i.e. *Oralment. Estudis de variació funcional* (1998) and *Corpus, corpora* (1996), respectively.

Bassols analysed Catalan riddles from the perspective of pragmatics (1990) and wrote an introductory text on pragmatics (2000). Castellà studied language usage (1992) and Lacreu became a best-selling author in Catalan linguistics with his *Manual d'ús de l'estàndard oral* published in 1990 (2002). All these scholars also played a highly significant role in applied linguistics as will be seen in the following section.

7.9 Applied linguistics: language teaching and learning, linguistic transference and the language of the media

Applied linguistics was promoted by M. Teresa Cabré (1990) and M. Josep Cuenca (1992), the latter did so by explaining how grammatical theories have been applied to language teaching. Payrató, after working on language interference between Spanish and Catalan (1985), once again lay the foundations for this promising new discipline (Payrató 1997).

With regard to L2 teaching in both senses, i.e. teaching Catalan as a foreign language and by teaching foreign languages to Catalan-speaking learners, Julià-Muné's contribution as an editor (2000a) is worth mentioning.¹⁸ As far as language usage in Catalan mass media is concerned, Francesc Vallverdú played an important role as a language advisor to the official Catalan radio and TV stations (see Appendix) and further elaborated his previous criteria (see 7.5) on the standard language designed for the oral media (2000). This issue had already been discussed at the University of Valencia (Ferrando 1990), at the IEC (1990a; 1996) and by university research groups through their publications in Catalonia (Bassols/Rico/Torrent 1997; Creus/Julià/Romero 2000; Cros/Segarra/Torrent 2000).

¹⁸ One of the most renowned linguists, Francesc de B. Moll, devoted himself to applied linguistics as a publisher of manuals for the study of German and French that were used in schools all over Spain for several decades after the Civil War. For details, see Julià-Muné (2004b).

7.10 The cradle of sociolinguistics

As it has been stated above (see 5), Professor Antoni M. Badia i Margarit tried to introduce what was supposed to be standard Catalan in his new grammar (1962) and elaborated it further in his widely circulated follow-up work (1964) and especially in his pioneering *La llengua dels barcelonins* five years later (1969). This last work came after he delivered a presentation on the sociolinguistic position of Catalan (Badia 1973c) at the first post-war congress on Catalan linguistics, held in Strasbourg in 1968 (see 6.1). Therefore, this promising new linguistic science can be said originated at the University of Barcelona, where Badia was lecturing. Sociolinguistics started being regularly taught at universities in 1976, precisely at the time when Catalan, as a result of the Spanish transition to democracy (1975–1982), was being institutionalised as an official language mainly in the areas of education and the mass media.

Then a lecturer was appointed to complement lectures on sociolinguistics at the University of Barcelona, Lluís V. Aracil (1941–), who was a challenging and outspoken intellectual always prepared to make his opinion known in a lucid and ironic manner. Apart from his lectures, he rose to prominence thanks to his pioneering and influential works (1966; 1975; 1982; 1983), such as *Dir la realitat*, in which he firmly stated what he believed the sociolinguist's job was:

‘Once personal and environmental differences are set aside, it can be seen that the sociolinguist is not an omniscient oracle nor a messianic titan. In spite of his work being indispensable – and I like to remind you that, according to [Joan] Fuster, *all* indispensable things are important – superhuman effort by itself is not sufficient at all. It seems clear to me that, in an ‘abnormal’ situation – which means a *very complicated* situation – effective clarification requires a *collective* effort. It is a matter of attention, imagination and reflection of many people – and of communication between them, of course – which must *create* the feeling that an otherwise solitary task is incapable of generating. You should think that this is really the *only* effective resource a society can mobilise to overcome confusion’ (Aracil 1983, 74).

Rafael L. Ninyoles (1943–), a Valencian like Aracil, further analysed the state of the language spoken in Valencia (1969) and later presented his views on the matter (1971) as well as on language policy (1976). The first two works were accompanied by prologues from Francesc Vallverdú and Joan Fuster, respectively. As well as being the most prolific author among them, the former also contributed to the dissemination of 20th century sociolinguistics in the Catalan-speaking territories, beginning with the results of languages in contact (Vallverdú 1968; 1970; 1980; 1990; 2000).

The pioneering work of these four sociolinguists culminated in 1998, when the Centre Universitari de Sociolingüística i Comunicació (CUSC) was founded at the University of Barcelona. It promotes interdisciplinary research and currently publishes the online journal *LSC – Llengua, Societat i Comunicació*.

7.11 Historiography of Catalan linguistics in the 20th and 21st centuries

Most Catalan linguists have written about their predecessors. I provide the following list of the most relevant works that reflects this interest: Alcover (1903; 1908c; 1909; 1915), Coromines (1943), Moll (²1981; ¹1962; 1983a; 1983b), Ferrater (1981; ¹1968), Argenter (1972; 2000), Murgades (1972), Rigau (1975), Badia i Margarit (1950; 1976; 1999; 2004), Gulsoy (1982), Segarra (1985a; 1985b; 1998), Colón/Soberanas (²1991; ¹1986), Bonet (2000), Colón (2001), Solà (1977a; 1987a; 1999), Rico/Solà (1995), Feliu (1997; 2000), Feliu/Albiol (2017); Ferrando/Pérez i Moragon (1998), Marcet/Solà (1998), Massot i Muntaner (1985; 1996; 2001), Veny (1999; 2002), Ginebra (1998; 2004), Cortés (2002), Iglésias (2004), Miralles (2003; 2005), Mascaró (1984; 2006), Perea (2003; 2005; 2006; see also Alcover 2003–2018, Moll 2003–2006), Ginebra/Solà (2007), Dols (2007), Ribes (2011), Mir/Solà (Fabra 2005–2013).

Nearing the end of the last great period of Catalan linguistics in the 20th century, there is one factor worth mentioning that affected its development. Since the 1980s it has been claimed for purely political reasons that Valencian is not just a geolectal variety of Catalan but a separate language. This led to the establishment of a new language academy in 1998, the Acadèmia Valenciana de la Llengua (see 8.2).

8 Contemporary linguistics II: new projects for the 21st century

Catalan linguistics has combined descriptive and prescriptive studies by editing and publishing two main grammars. Among the descriptive studies the *Gramàtica del català contemporani* (GCC, ¹2002; ⁴2008) is worthy of particular attention. It was conceived and directed by Joan Solà and co-edited by Maria-Rosa Lloret (phonetics and phonology), Joan Mascaró (morphology) and Manuel Pérez Saldanya (syntax). Volume 1 covers phonetics and phonology (11 chapters) as well as morphology (10 chapters); volumes 2 and 3 contain 31 chapters dealing with syntax. Among the so-called prescriptive studies the *Gramàtica de la llengua catalana* (IEC 2016), stands out in particular. Solà again masterminded its chapters on syntax and worked directly on them, assisted by Gemma Rigau and Pérez Saldanya, until his death in 2010.

8.1 The Institut d'Estudis Catalans and its works on linguistics

The prescriptive work par excellence updates, as we have just seen, the official normative grammar (Fabra ¹1918, ⁷1933). The new *Ortografia catalana* (2017) was

published separately. In terms of lexicography, the second edition of the prescriptive dictionary of the Institute, known as *DIEC2*,¹⁹ was published in 2007.

Over the last three decades the IEC has been developing the programme *Diccionari del català contemporani* (DCC). In its pre-lexicographic phase the *Corpus Textual Informatitzat de la Llengua Catalana* (CTILC, 79 Language Corpora) was compiled and completed (1985–1997). The output of the CTILC includes a frequency dictionary (Rafel i Fontanals 1996a; 1998a; 1998b). The lexicographic phase of DCC is the *Diccionari descriptiu de la llengua catalana* (<http://dcc.iec.cat/ddlci/scripts/index1.asp>), based on CTILC.

In the field of onomastics the *Nomenclàtor* (2009; 2003) has been revised and extended. It provides 52,000 entries (13,000 more than the first edition) and the phonetic transcription for almost 1,000 toponyms. Dialectological and geolinguistic works include the *Atles Lingüístic del Domini Català*, with 9 volumes covering a wide range of topics, starting with ‘Human body and diseases’ (Veny/Pons 2001–2017), and its sequel, based on the maps of the above-mentioned atlas, the *Petit Atles Lingüístic del Domini Català* (2007–), which is still in progress.

In phonetics Daniel Recasens’s works include the results of his research in the IEC’s Laboratory (2004; 2011; 2014). He takes advantage of the production and perception phenomena to explain phonetic changes in the various varieties of Catalan, which has been widely applied to his recent *Fonètica històrica del català* (2017). Recasens synthesises and enhances the work of a whole century in both experimental and historical phonetics.

8.2 Contributions of the Acadèmia Valenciana de la Llengua (AVL)

The Valencian Language Academy has been engaged in a wide range of linguistic activities and been producing accompanying publications for the last two decades since its foundation in 1998. The following works, presented in chronological order, are mentioned based on their relevance to the codification and standardisation of the Valencian variety of Catalan (8.2 Social and Functional Variation in Catalan): *Gramàtica normativa valenciana* (AVL 2006a), *Diccionari ortogràfic i de pronunciació del valencià* (AVL 2006b) and the practical and complementary *La normativa ortogràfica del valencià* (AVL 2006c), integrated online with the *Gramàtica normativa; L’estàndard oral valencià* (AVL 2008), *Corpus toponímic valencià* (AVL 2009), and more recently, *Diccionari normatiu valencià* (AVL 2016a), as well as popularised versions such as *Gramàtica valenciana bàsica* (AVL 2016b).

¹⁹ The Institute has new tools at its disposal to take advantage of the best online ways to access the results of all its available dictionaries.

8.3 Work in progress: popularising linguistics, lexicography, language usage, oral standardisation, neology, linguistic geography, onomastics and applied linguistics

The new century saw the publication of the *Enciclopèdia de la llengua catalana* (ELC, 2001), edited by Vallverdú with the help of 88 contributors. It was inspired in particular by the *Cambridge Encyclopedia of the English Language* (Crystal 1995). In onomastics contributions were made by Moran/Batlle/Rabella (2002) and a new journal *Onomàstica* was issued in 2015. In phonology Wheeler updated his 1979 work (2005), in discourse analysis a notable contribution was made by Castellà (2004), in applied phonetics by Julià-Muné (2005c), in language history by Rasico (2006), in elocution and orthology by Oliva (2006) and Rossich (2006), respectively, in pragmatics (Bassols 2007), in oral usage in the media (Julià-Muné 2004a), in morphosyntax the dictionary of verbal usage was updated (Ginebra/Montserrat 2009) and in intonation Prieto/Cabré (2013) was also greeted with acclaim. Other lexicographical projects have proved fruitful such as those related to neology (Cabré/Domènech/Estopà 2014; Freixa/Bernal/Cabré 2015), promoted by M. Teresa Cabré as current President of the Philology Section of the IEC, where linguistic research is undertaken and shared on a wide range of subjects, encompassing grammar, lexicon, terms, the oral standard, dialects, onomastics and phonetics.

The Catalan Academy continues living up to the standards devised by its founder, Enric Prat de la Riba, at the centenary of his death. Fortunately, the Academy's enthusiasm, competence and effectiveness are matched elsewhere in the country, mainly through university research projects undertaken by a community of academics working in areas ranging from sociolinguistics, geolinguistics, pragmatics, syntax, phonology and morphology, L2 pronunciation teaching, theoretical linguistics, language assessment, translation and interpretation, speech therapy, etc.

At this point we may conclude with a reminder that the history of words and grammars of languages will continue to be written as long as mankind makes use of speech in order to communicate effectively through both images and sound. About three decades after the Prussian philosopher and linguist Wilhelm von Humboldt's holistic education model reached the three young Catalan scholarship holders at Halle University, and three decades before Noam Chomsky's ideas about language became known worldwide, William Somerset Maugham put it this way in his literary memoir *The Summing Up*: "Words have weight, sound and appearance; it is only by considering these that you can write a sentence that is good to look at and good to listen to. [...] [After all] it is well to remember that grammar is common speech formulated" (1938, 13).

9 Bibliography

- Aguiló, Marià (1915–1934), *Diccionari Aguiló; materials lexicogràfics aplegats per Marian Aguiló i Fuster, revisats i publicats sota la cura de Pompeu Fabra i Manuel de Montoliu*, 8 vol., Barcelona, Institut d'Estudis Catalans.
- Alarcos Llorach, Emilio (1983), *Estudis de lingüística catalana*, Barcelona, Ariel.
- Albaigès, Josep M. (¹12000; ¹1980), *Diccionari de noms de persona*, Barcelona, Edicions 62.
- Alcover, Antoni M. (²1902a; ¹1901), *Diccionari de la llengua catalana. Lletre de convit que a tots els amichs d'aquesta llengua envia Mossen Antoni M.^a Alcover, Pre Vicari General de Mallorca*, Barcelona, “La Catalana” de J. Puigventós (1902). New updated edition by Aina Moll (2003), Palma de Mallorca, Moll.
- Alcover, Antoni M. (1902b), *Les paraules s'han de replegar tal com el poble les pronuncia. Trascendencia capital de la pronunciació popular*, Bolletí del Diccionari de la Llengua Catalana 1, 139–142.
- Alcover, Antoni M. (1903), *Questions de llengua y literatura catalana*, Bolletí del Diccionari de la Llengua Catalana 1, 209–560.
- Alcover, Antoni M. (1907), *Dietari de l'escursió filològica feta ab el Dr. Schädel, de cap a cap dels Pirineus catalans (31 juliol–13 setembre [1906])*, Bolletí del Diccionari de la Llengua Catalana 3, 257–367.
- Alcover, Antoni M. (1908a), *Concordansa del participi ab el terme d'acció*, in: *Primer Congrés Internacional de la Llengua Catalana. Barcelona. Octubre de 1906*, Barcelona, Joaquim Horta, 124–128.
- Alcover, Antoni M. (1908b), *La llengua catalana té sintaxis pròpia*, in: *Primer Congrés Internacional de la Llengua Catalana. Barcelona. Octubre de 1906*, Barcelona, Joaquim Horta, 350–399.
- Alcover, Antoni M. (1908c), *El Dr. Milà i Fontanals i la filologia catalana*, Bolletí del Diccionari de la Llengua Catalana 4, 75–79.
- Alcover, Antoni M. (1909), *Una mica de Dialectologia catalana. L'obra de Mr. Morel-Fatio i Mr. Saroihandy dins “Grundriss” del Dr. Gröber sobre el català*, Bolletí del Diccionari de la Llengua Catalana 4, 194–303.
- Alcover, Antoni M. (1913), *Dietari de l'eixida d'enguany a França i altres nacions*, Bolletí del Diccionari de la Llengua Catalana 7, 309–356.
- Alcover, Antoni M. (1915), *Pròleg*, in: Tomàs Forteza, (1892–1898/1915), *Gramàtica de la llengua catalana*, [Palma de Mallorca], Escuela Tipográfica Provincial, 5–87.
- Alcover, Antoni M. (1922), *Dietari de l'Eixida Filològica per tot Catalunya i el Reyne de València de dia 26 de juny fins a 23 de novembre de 1921*, Bolletí del Diccionari de la Llengua Catalana 12 (1921–1922), 145–335.
- Alcover, Antoni M. (2003–2018), *Obres completes*, ed. M. Pilar Perea, 4 vol., Palma de Mallorca, Moll.
- Alcover, Antoni M. (2004, ¹1908–1926), *La pronúncia llatina entre catalans (Estudis fonològics)*, ed. Nicolau Dols, Barcelona/Palma, Publicacions de l'Abadia de Montserrat/Universitat de les Illes Balears.
- ALPI = Navarro Tomás, Tomás (ed.) (1962), *Atlas Lingüístico de la Península Ibérica*, Madrid, CSIC.
- Aracil, Lluís V. (1966), *Un dilema valencià/A Valencian dilemma*, Identity Magazine 24, 17–29.
- Aracil, Lluís V. (1975), *La revolució sociolingüística catalana*, Presència (13th Dec.).
- Aracil, Lluís V. (1983), *Dir la realitat*, Barcelona, Països Catalans.
- Aracil, Lluís V. (²1986, ¹1982), *Papers de sociolingüística*, Barcelona, La Magrana.
- Aragonès, Albert (1995), *La llengua del Baix Ebre i del Montsià. Un model de llengua estàndard oral*, Tortosa, Direcció General de Política Lingüística.
- Aramon i Serra, Ramon (ed.) (²1971, ¹1963–1968), *Estudis de lingüística i filologia catalanes dedicats a la memòria de Pompeu Fabra en el centenari de la seva naixença*, 2 vol., Estudis Romànics 22/23, Barcelona, Institut d'Estudis Catalans.

- Aramon i Serra, Ramon (1973), *Problèmes d'histoire de la langue catalane*, in: Antoni M. Badia i Margarit/Georges Straka (edd.), *La linguistique catalane. Colloque International organisé par le Centre de Philologie et Littératures Romanes de Strasbourg (1968)*, Paris, Klincksieck, 27–70.
- Argenter, Joan A. (1972), *[Gabriel Ferrater 1922–1972] Comentaris a la seva obra lingüística*, Serra d'Or 24, 395–396.
- Argenter, Joan A. (ed.) (2000), *Simposi Pompeu Fabra*, Barcelona, Institut d'Estudis Catalans.
- Arteaga, Josep M. (1904), *Spécimen catalan d'après la prononciation normale*, Le Maître Phonétique 19, 118–123.
- Arteaga, Josep M. (1908), *Ullada general a la fonètica catalana. El seu caràcter propi dins la família novo-llatina*, in: *Primer Congrès Internacional de la Llengua Catalana. Barcelona. Octubre de 1906*, Barcelona, Joaquim Horta, 445–465.
- Arteaga, Josep M. (1915), *Textes catalans avec leur transcription phonétique précédés d'un aperçu sur les sons du catalan*, ed. Pere Barnils, Barcelona, Institut d'Estudis Catalans.
- AVL (2006a) = Acadèmia Valenciana de la Llengua, *Gramàtica normativa valenciana*, València, Acadèmia Valenciana de la Llengua.
- AVL (2006b) = Acadèmia Valenciana de la Llengua, *Diccionari ortogràfic i de pronunciació del valencià*, València, Acadèmia Valenciana de la Llengua.
- AVL (2006c) = Acadèmia Valenciana de la Llengua, *La normativa ortogràfica del valencià*, www.avl.gva.es/documents/35882/40728/Ortograf.pdf/5a4e9e3e-b04e-43f0-85d1-61debac551ff (last accessed 24.02.2018).
- AVL (2008) = Acadèmia Valenciana de la Llengua, *L'estàndard oral valencià*, www.avl.gva.es/documents/35882/40728/Oral.pdf/97c04880-4477-48b2-a314-00e626f43674 (last accessed 24.02.2018).
- AVL (2009) = Acadèmia Valenciana de la Llengua, *Corpus toponímic valencià*, València, Acadèmia Valenciana de la Llengua.
- AVL (2016a) = Acadèmia Valenciana de la Llengua, *Diccionari normatiu valencià*, 2 vol., València, Acadèmia Valenciana de la Llengua.
- AVL (2016b) = Acadèmia Valenciana de la Llengua, *Gramàtica valenciana bàsica*, València, Acadèmia Valenciana de la Llengua.
- Badia i Margarit, Antoni M. (1950), “Regles d'esquivar vocables o mots grossers o pagesívols”. *Unas normas del s. XV sobre pureza de la lengua catalana, I: Edición del texto*, Boletín de la Real Academia de Buenas Letras de Barcelona 23, 137–152.
- Badia i Margarit, Antoni M. (1951), *Gramàtica històrica catalana*, Barcelona, Noguer; Catalan version: (1981), *Gramàtica històrica catalana*, València, 314.
- Badia i Margarit, Antoni M. (1962), *Gramàtica catalana*, 2 vol., Madrid, Gredos.
- Badia i Margarit, Antoni M. (1964), *Llengua i cultura als Països Catalans*, Barcelona, Edicions 62.
- Badia i Margarit, Antoni M. (1969), *La llengua dels barcelonins. Resultats d'una enquesta sociològico-lingüística*, Barcelona, Edicions 62.
- Badia i Margarit, Antoni M. (1973a), *Phonétique et phonologie catalanes*, in: Antoni M. Badia i Margarit/Georges Straka (edd.), *La linguistique catalane. Colloque International organisé par le Centre de Philologie et Littératures Romanes de Strasbourg (1968)*, Paris, Klincksieck, 115–166.
- Badia i Margarit, Antoni M. (1973b), *Morphosyntaxe catalane*, in: Antoni M. Badia i Margarit/Georges Straka (edd.), *La linguistique catalane. Colloque International organisé par le Centre de Philologie et Littératures Romanes de Strasbourg (1968)*, Paris, Klincksieck, 181–229.
- Badia i Margarit, Antoni M. (1973c), *Le catalan aujourd'hui*, in: Antoni M. Badia i Margarit/Straka, Georges (edd.), *La linguistique catalane. Colloque International organisé par le Centre de Philologie et Littératures Romanes de Strasbourg (1968)*, Paris, Klincksieck, 379–443.
- Badia i Margarit, Antoni M. (1976), *Vint-i-cinc anys d'estudis sobre la llengua i la literatura catalanes (1950–1975). I: La llengua*, Barcelona, Publicacions de l'Abadia de Montserrat.

- Badia i Margarit, Antoni M. (1977), *Ciència i passió dins la cultura catalana*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Badia i Margarit, Antoni M. (1988), *Sons i fonemes de la llengua catalana*, Barcelona, Universitat de Barcelona.
- Badia i Margarit, Antoni M. (1994), *Gramàtica de la llengua catalana. Descriptiva, normativa, diatòpica, diastràtica*, Barcelona, Enciclopèdia Catalana.
- Badia i Margarit, Antoni M. (ed.) (1999), *Les "Regles de esquivar vocables" i "La qüestió de la llengua"*, Barcelona, Institut d'Estudis Catalans.
- Badia i Margarit, Antoni M. (2004), *Moments clau de la història de la llengua catalana*, ed. Antoni Ferrando, València, Universitat de València.
- Badia i Margarit, Antoni M./Straka, Georges (edd.) (1973), *La linguistique catalane. Colloque International organisé par le Centre de Philologie et Littératures Romanes de Strasbourg (1968)*, Paris, Klincksieck.
- Balari i Jovany, Josep (1926–1936), ed. Manuel de Montoliu, 2 vol. uncompleted, Barcelona, Elzeviriana/Camí.
- Ballot i Torres, Josep Pau (1815), *Gramatica y apologia de la llengua cathalana*, Barcelona, Juan Francisco Piferrer.
- Balsalobre, Pep/Gratacós, Joan (edd.) (1995), *La llengua catalana al segle XVIII*, Barcelona, Quaderns Crema.
- Barnils, Pere (1912), *Études de prononciations catalanes à l'aide du palais artificiel*, Revue de Phonétique 2, 50–68.
- Barnils, Pere (ed.) (1917), *Estudis fonètics*, Barcelona, Laboratori de Fonètica Experimental de l'Institut d'Estudis Catalans.
- Barnils, Pere (1918), *Treballs realitzats durant l'any 1917*, Barcelona, Laboratori de Fonètica Experimental de l'Institut d'Estudis Catalans.
- Barnils, Pere (ed.) (1918–1921), *La Paraula*, Barcelona, Escola Municipal de Sords-muts.
- Barnils, Pere (1930), *Defectes del parlar*, Barcelona, Catalònia.
- Barnils, Pere (ed.) (1931–1932), *El Parlar. Butlletí que publica el Laboratori de la Paraula*, Barcelona, Laboratori de la Paraula.
- Bassols, Margarida (1990), *Anàlisi pragmàtica de les endevinalles catalanes*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Bassols, Margarida (2000), *Les claus de la pragmàtica*, Vic, Eumo.
- Bassols, Margarida (2007), *El llenguatge dels polítics*, Vic, Eumo.
- Bassols, Margarida/Rico, Albert/Torrent, Anna M. (edd.) (1997), *La llengua de TV3*, Barcelona, Empúries.
- BDLC = *Bolletí del Diccionari de la Llengua Catalana*, Palma de Mallorca, ed. Antoni M. Alcover (1901–1926) and Francesc de B. Moll (1933–1937).
- Bofarull, Antoni de (1864), *Estudios, sistema gramatical y crestomatía de la lengua catalana*, Barcelona, Impr. de Lluís Tasso.
- Bonet, Eulàlia/Lloret, Maria-Rosa (1997), *Fonologia catalana*, Barcelona, Ariel.
- Bonet, Eulàlia/Lloret, Maria-Rosa/Mascaró, Joan (2000; 1997), *Manual de transcripció fonètica*, Bellaterra, Universitat Autònoma de Barcelona.
- Bonet, Sebastià (2000), *Les gramàtiques normatives valencianes i balears del segle XX*, València, Universitat de València.
- Bonet, Sebastià/Solà, Joan (1986), *Sintaxi generativa catalana*, Barcelona, Enciclopèdia Catalana.
- Bover, August (1993), *Manual de catalanística*, Tarragona, Diputació de Tarragona/Barcelona, Publicacions de l'Abadia de Montserrat.
- Bruguera, Jordi (1985), *Història del lèxic català*, Barcelona, Enciclopèdia Catalana.
- Bruguera, Jordi (2004; 1990), *Diccionari ortogràfic i de pronúncia*, Barcelona, Enciclopèdia Catalana.

- Butlletí de Dialectologia Catalana (1913–1937), Barcelona, Institut d'Estudis Catalans, edited successively by Pere Barnils, Antoni Griera, Pompeu Fabra and Joan Coromines.
- Cabré, M. Teresa (ed.) (1990), *La lingüística aplicada*, Barcelona, “la Caixa”/Universitat de Barcelona.
- Cabré, M. Teresa (1992), *La terminologia. La teoria, el mètode, les aplicacions*, Barcelona, Empúries.
- Cabré, M. Teresa/Domènech, Ona/Estopà, Rosa (edd.) (2014), *Mots nous en català. New Words in Catalan*, Amsterdam, Benjamins.
- Cabré, M. Teresa/Rigau, Gemma (1985), *Lexicologia i semàntica*, Barcelona, Enciclopèdia Catalana.
- Cabré, M. Teresa, et al. (edd.) (1984), *Problemàtica de la Normativa del Català. Actes de les Primeres Jornades d'Estudi de la Llengua Normativa*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Cabré, M. Teresa, et al. (edd.) (1987), *Actes de les Segones Jornades d'Estudi de la Llengua Normativa*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Calaforra, Guillem (1998), *Wilhelm Meyer-Lübke i “Das Katalanische”. Introducció i traducció*, Barcelona, Institut d'Estudis Catalans.
- Calveras, Josep (1925), *La reconstrucció del llenguatge literari català. Estudi d'orientació*, Barcelona, Biblioteca Balmes.
- Cardellach, Pau (2017, ¹1840), *Gramàtica catalana*, ed. Olga Fullana, Barcelona, Institut d'Estudis Catalans.
- Castellà, Josep M. (1992), *De la frase al text. Teories de l'ús lingüístic*, Barcelona, Empúries.
- Castellà, Josep M. (2004), *Oralitat i escriptura. Dues cares de la complexitat del llenguatge*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Cerdà, Ramon (1972), *El timbre vocàlic en catalán*, Madrid, CSIC.
- CILC (1908), *Primer Congrés Internacional de la Llengua Catalana. Barcelona. Octubre de 1906*, Barcelona, Joaquim Horta.
- Colomina, Jordi (1985), *L'alacantí. Un estudi sobre la variació lingüística*, Alacant, Institut d'Estudis “Juan Gil-Albert”.
- Colomina, Jordi (1999), *Dialectologia catalana. Introducció i guia bibliogràfica*, Alacant, Secretariat de Publicacions i Departament de Filologia Catalana de la Universitat d'Alacant.
- Colón, Germà (1973), *Quelques considérations sur le lexique catalan*, in: Antoni M. Badia i Margarit/Georges Straka (edd.), *La linguistique catalane. Colloque International organisé par le Centre de Philologie et Littératures Romanes de Strasbourg* (1968), Paris, Klincksieck, 240–280.
- Colón, Germán (1976), *El léxico catalán en la Romania*, Madrid, Gredos. Catalan translation: (1993), *El lèxic català dins la Romània*, València, Universitat de València.
- Colón, Germà (1978), *La llengua catalana en els seus textos*, 2 vol., Barcelona, Curial.
- Colón, Germà (1989), *El español y el catalán, juntos y en contraste*, Barcelona, Ariel.
- Colón, Germà (2001), *Les “Regles d'esquivar vocables”. Autoria i entorn lingüístic*, Barcelona, Institut d'Estudis Catalans.
- Colón, Germà/Soberanas, Amadeu-J. (²1991, ¹1986), *Panorama de la lexicografia catalana. De les glosses medievals a Pompeu Fabra*, Barcelona, Enciclopèdia Catalana.
- Coromines, Joan (1936), *El parlar de Cardós i Vall Ferrera*, Butlletí de Dialectologia Catalana 23, 1–81; (²1976), also in: Joan Coromines (1976–1977), *Entre dos llenguatges*, Barcelona, Curial, vol. 2, 29–67.
- Coromines, Joan (1937), *Mots catalans d'origen aràbic*, Butlletí de Dialectologia Catalana 24, 1–81; (²1977), also in: Joan Coromines (1976–1977), *Entre dos llenguatges*, Barcelona, Curial, vol. 3, 68–177.
- Coromines, Joan (ed.) (1943), *Miscel·lània Fabra. Recull de treballs de lingüística catalana i romànica dedicats a Pompeu Fabra pels seus amics i deixebles amb motiu del 75è aniversari de la seva naixença*, Buenos Aires, “Coni”.
- Coromines, Joan (¹1958), *De gramàtica històrica catalana: a propòsit de dos llibres*, in: Anna Granville Hatcher/Karl Ludwig Selig (edd.), *Studia Philologica et Litteraria in Honorem Leo Spitzer*, Bern,

- Francke, 123–148; also in: Joan Coromines (²1974, 1971b), *Lleures i converses d'un filòleg*, Barcelona, Club Editor, 245–275.
- Coromines, Joan (¹1963), *Sobre l'elocució catalana en el teatre i en la recitació*, in: *Miscel·lània Carles Cardó*, Barcelona, Ariel, 339–348; also in: Joan Coromines (²1974, 1971b), *Lleures i converses d'un filòleg*, Barcelona, Club Editor, 94–105.
- Coromines, Joan (1965–1970), *Estudis de toponímia catalana*, 2 vol., Barcelona, Barcino.
- Coromines, Joan (1971a), *Alguns noms propis estrangers usats sovint en forma incorrecta*, in: Joan Coromines, *Lleures i converses d'un filòleg*, Barcelona, Club Editor, 55–60.
- Coromines, Joan (²1974, 1971b), *Lleures i converses d'un filòleg*, Barcelona, Club Editor.
- Coromines, Joan (1976–1977), *Entre dos llenguatges*, 3 vol., Barcelona, Curial.
- Coromines, Joan (1980–2001), *Diccionari etimològic i complementari de la llengua catalana*, 10 vol., Barcelona, Curial.
- Coromines, Joan (1989–1997), *Onomasticon Cataloniae*, 8 vol., Barcelona, Curial.
- Cortés, Santi (2002), *Manuel Sanchis Guarner (1911–1981). Una vida per al diàleg*, València/Barcelona, Institut Interuniversitari de Filologia Valenciana/Publicacions de l'Abadia de Montserrat.
- Creus, Imma/Julià, Joan/Romero, Sílvia (edd.) (2000), *Llengua i mitjans de comunicació. Actes del Congrés de Llengua i Mitjans de Comunicació*, Lleida, Pagès.
- Cros, Anna/Segarra, Mila/Torrent, Anna M. (edd.) (2000), *Llengua oral i llengua escrita a la televisió*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Crystal, David (ed.) (1995), *Cambridge Encyclopedia of the English Language*, Cambridge, Cambridge University Press.
- Cuenca, M. Josep (1988–1991), *L'oració composta*, 2 vol., València, Universitat de València.
- Cuenca, M. Josep (1992), *Teories gramaticals i ensenyament de llengües*, València, Tàndem.
- DCVB = Antoni M. Alcover/Francesc de B. Moll (1926–1962), *Diccionari català-valencià-balear*, 10 vol., Palma de Mallorca, Moll.
- DGLC = Fabra, Pompeu (1932b), *Diccionari general de la llengua catalana*, Barcelona, Catalonia.
- Diputació de Barcelona (1916), *Guia de les institucions científiques i d'ensenyança*, Barcelona, Consell de Pedagogia.
- Dols, Nicolau (2007), *Els estudis de fonologia catalana en els darrers cent anys*, in: *Actes del Tretzè Col·loqui de l'AILLC*, vol. 1, Barcelona, Publicacions de l'Abadia de Montserrat, 323–359.
- Duarte, Carles/Alsina, Àlex (1984), *Gramàtica històrica del català*, 3 vol., Barcelona, Curial.
- ELC = Francesc Vallverdú (ed.) (2001), *Enciclopèdia de la llengua catalana*, Barcelona, Edicions 62.
- Fabra, Pompeu (1891), *Ensayo de gramática de catalán moderno*, Barcelona, L'Avenç.
- Fabra, Pompeu (1897), *Étude de phonologie catalane (catalan oriental)*, *Revue Hispanique* 4, 5–30.
- Fabra, Pompeu (1898), *Contribució a la gramàtica de la llengua catalana*, Barcelona, L'Avenç.
- Fabra, Pompeu (1903), *La representació de la "l" palatal en català*, *Bolletí del Diccionari de la Llengua Catalana* 1, 562–569.
- Fabra, Pompeu (1905), *Les tres preteses lleis d'escursament, de distinció y d'eufonia*, *El Poble Català* (9th, 16th September).
- Fabra, Pompeu (1906), *Les "e" toniques du catalan*, *Revue Hispanique* 15, 9–23.
- Fabra, Pompeu (1907), *Le catalan dans la "Grammaire des langues romanes" de W. Meyer-Lübke et dans le "Grundriss der rom. Philologie"*, *Revue Hispanique* 17, 1–45.
- Fabra, Pompeu (1908), *Qüestions d'ortografia catalana*, in: *Primer Congrés Internacional de la Llengua Catalana. Barcelona. Octubre de 1906*, Barcelona, Joaquim Horta, 188–221.
- Fabra, Pompeu (1912), *Gramàtica de la llengua catalana*, Barcelona, L'Avenç.
- Fabra, Pompeu (1913), *Els mots àtons en el parlar de Barcelona*, I, *Butlletí de Dialectologia Catalana* 1, 7–17.
- Fabra, Pompeu (1914), *Els mots àtons en el parlar de Barcelona*, II, *Butlletí de Dialectologia Catalana* 2, 1–6.

- Fabra, Pompeu (1926), *La coordinació i la subordinació en els documents de la Cancilleria catalana durant el segle XIVè*, in: *Memòries de la Secció Filològica de l'Institut d'Estudis Catalans*, vol. 1/1, Barcelona, Institut d'Estudis Catalans, 1–30.
- Fabra, Pompeu (1932a), *El català literari*, Barcelona, Barcino.
- Fabra, Pompeu (¹1932b), *Diccionari general de la llengua catalana*, Barcelona, Catalonia.
- Fabra, Pompeu (⁷1933, ¹1918), *Gramàtica catalana*, Barcelona, Institut d'Estudis Catalans.
- Fabra, Pompeu (⁴1937, ³1931, ²1923, ¹1917), *Diccionari ortogràfic*, Barcelona, Institut d'Estudis Catalans.
- Fabra, Pompeu (1941), *Grammaire catalane*, Paris, Les Belles Lettres.
- Fabra, Pompeu (1946), *Gramàtica catalana*, París, Cultura Catalana.
- Fabra, Pompeu (²1954), *Diccionari general de la llengua catalana*, Barcelona, A. López Llausàs.
- Fabra, Pompeu (1954–1956, ¹1919–1928), *Converses filològiques*, ed. Santiago Pey, 10 vol., Barcelona, Barcino.
- Fabra, Pompeu (1956), *Gramàtica catalana*, ed. Joan Coromines, Barcelona, Teide.
- Fabra, Pompeu (1983–1984, ¹1919–1928), *Converses filològiques*, ed. Joaquim Rafel, 2 vol., Barcelona, Edhasa.
- Fabra, Pompeu (2005–2013), *Obres completes*, 9 vol., edd. Jordi Mir/Joan Solà, Barcelona/València/Palma, Proa/3i4/Moll.
- [Fabra, Pompeu/Barnils, Pere] (1915), *Report dels treballs fets per les Oficines lexicogràfiques durant el bienni de 1913–1914*, Barcelona, L'Avenç.
- [Fabra, Pompeu/Barnils, Pere] (1917), *Report dels treballs fets per les Oficines lexicogràfiques durant el bienni de 1915–1916*, Barcelona, L'Avenç.
- Fabra, Pompeu/Montoliu, Manuel (edd.) (1915–1934), *Diccionari Aguiló*, 8 vol., Barcelona, Institut d'Estudis Catalans.
- Febrer i Cardona, Antoni (2004), *Obres gramaticals I*, ed. Jordi Ginebra, Barcelona, Institut d'Estudis Catalans.
- Febrer i Cardona, Antoni (2017), *Obres gramaticals II*, edd. Francesc Feliu/Gemma Albiol, Barcelona, Institut d'Estudis Catalans.
- Feliu, Francesc (1997), *L'obra filològica d'Antoni de Bastero i Lledó. Edició de la "Història de la llengua catalana"*, PhD thesis, Girona, Universitat de Girona.
- Feliu, Francesc (2000), *Catàleg dels manuscrits filològics d'Antoni de Bastero*, Barcelona, Institut d'Estudis Catalans.
- Feliu, Francesc/Albiol, Gemma (2017), *Estudi introductori*, in: Antoni Febrer i Cardona, *Obres gramaticals II*, edd. Francesc Feliu/Gemma Albiol, Barcelona, Institut d'Estudis Catalans, 9–40.
- Ferrando, Antoni (ed.) (1990), *La llengua als mitjans de comunicació. Actes de les Jornades sobre la Llengua Oral als Mitjans de Comunicació Valencians*, València, Institut de Filologia Valenciana, Universitat de València.
- Ferrando, Antoni/Nicolàs, Miquel (2005), *Història de la llengua catalana*, Barcelona, Pòrtic/UOC.
- Ferrando, Antoni/Pérez i Moragon, Francesc (edd.) (1998), *Manuel Sanchis Guarner: el compromís cívic d'un filòleg*, València, Universitat de València.
- Ferrater, Gabriel (1981, ¹1968), *Les gramàtiques de Pompeu Fabra*, in: *Sobre el llenguatge*, ed. Joan Ferraté, Barcelona, Quaderns Crema, 3–12.
- Forteza, Tomàs (1915), *Gramàtica de la llengua catalana*, ed. Antoni M. Alcover, [Palma de Mallorca], Escuela Tipográfica Provincial.
- Freixa, Judit/Bernal, Elisenda/Cabrè, M. Teresa (edd.) (2015), *La neologia lèxica catalana*, Barcelona, Institut d'Estudis Catalans.
- Galí, Alexandre (1926), *L'ensenyament de l'ortografia als infants*, Barcelona, Barcino.
- Galí, Alexandre (1931a), *Lliçons de llenguatge*, Barcelona, Pedagògica.

- Galí, Alexandre (1931b), *Per la llengua i per l'escola*, Barcelona, La Revista.
- Galí, Alexandre (1935), *Introducció a la gramàtica*, Barcelona, Proa.
- GCC = Joan Solà, et al. (edd.) (2008, 2002), *Gramàtica del català contemporani*, 3 vol., Barcelona, Empúries.
- GIEC = Institut d'Estudis Catalans (2016), *Gramàtica de la llengua catalana*, Barcelona, Institut d'Estudis Catalans.
- Gili, Joan (1974, 1943), *Introductory Catalan Grammar*, Oxford, The Dolphin Book.
- Ginebra, Jordi (1998), *Estudi introductori i edició*, in: Joan Petit i Aguilar, *Gramàtica catalana*, Barcelona, Institut d'Estudis Catalans, 15–105.
- Ginebra, Jordi (2004), *Estudi introductori i edició*, in: Antoni Febrer i Cardona, *Obres gramaticals I*, Barcelona, Institut d'Estudis Catalans, 9–36.
- Ginebra, Jordi/Montserrat, Anna (2009, 1999), *Diccionari d'ús dels verbs catalans*, Barcelona, Edicions 62.
- Ginebra, Jordi/Solà, Joan (2007), *Pompeu Fabra: vida i obra*, Barcelona, Teide.
- Gröber, Gustav (ed.) (1904-1906, 1888–1897), *Grundriss der romanischen Philologie*, 2 vol., Strassburg, Trübner.
- Guiter, Henri (1973), *Questions d'onomastique catalane*, in: Antoni M. Badia i Margarit/Georges Straka (edd.), *La linguistique catalane. Colloque International organisé par le Centre de Philologie et Littératures Romanes de Strasbourg (1968)*, Paris, Klincksieck, 339–367.
- Gulsoy, Joseph (1982), *Catalan*, in: Rebecca Posner/John N. Green, (edd.), *Language and Philology in Romance*, The Hague, Mouton, 189–296.
- Gulsoy, Joseph (1993), *Estudis de gramàtica històrica*, València/Barcelona, Institut Interuniversitari de Filologia Valenciana/Publicacions de l'Abadia de Montserrat.
- Hoenigswald, Henry M. (1986), *Nineteenth-century linguistics on itself*, in: Theodora Bynon/Frank Palmer (edd.), *Studies in the history of Western linguistics in Honour of R. H. Robins*, Cambridge, Cambridge University Press, 172–188.
- Hualde, José Ignacio (1992), *Catalan*, London/New York, Routledge.
- Iglésias, Narcís (2004), *Una revisió de Fabra, una crítica a la norma. L'obra lingüística de Josep Calveras*, Girona, CCG/UdG.
- Institut d'Estudis Catalans (1913), *Normes ortogràfiques*, Barcelona, Institut d'Estudis Catalans.
- Institut d'Estudis Catalans (1990a, 1986), *La llengua en els mitjans de comunicació de massa i especialment en la televisió*, reprinted in: *Documents de la Secció Filològica I*, Barcelona, Institut d'Estudis Catalans, 137–140.
- Institut d'Estudis Catalans (1990b), *Proposta per a un estàndard oral*, vol. I: *Fonètica*, ed. Joan A. Argenter, Barcelona, Institut d'Estudis Catalans.
- Institut d'Estudis Catalans (1992), *Proposta per a un estàndard oral*, vol. II: *Morfologia*, ed. Joan A. Argenter, Barcelona, Institut d'Estudis Catalans.
- Institut d'Estudis Catalans (2007, 1995), *Diccionari de la llengua catalana*, Barcelona/València/Palma de Mallorca, Enciclopèdia Catalana – Edicions 62/3 i 4/Moll.
- Institut d'Estudis Catalans (1996, 1991), *La llengua i els mitjans de comunicació de massa*, reprinted in: *Documents de la Secció Filològica II*, Barcelona, Institut d'Estudis Catalans, 31–36.
- Institut d'Estudis Catalans (2009, 2003), *Nomenclàtor oficial de toponímia major de Catalunya*, Barcelona, Generalitat de Catalunya/Institut d'Estudis Catalans.
- Institut d'Estudis Catalans (2017), *Ortografia catalana*, Barcelona, Institut d'Estudis Catalans.
- Julià i Muné, Joan (1984), *The Birth of Experimental Phonetics in Spain. Pere Barnils and his Barcelona Laboratories (1913–1933)*, in: Marcel P. R. Van den Broecke/Antonie Cohen (edd.), *Proceedings of the 10th International Congress of Phonetic Sciences (Utrecht, 1983)*, Dordrecht, Foris, 667–669.

- Julià i Muné, Joan (ed.) (2000a), *L'ensenyament del català com a L2. De la teoria a la pràctica*, Lleida, Universitat de Lleida.
- Julià i Muné, Joan (2000b), *L'inici de la lingüística catalana: Bernhard Schädcl, Mn. Antoni M. Alcover i l'Institut d'Estudis Catalans. Una aproximació epistolar (1904–1925)*, Barcelona, Curial/Publicacions de l'Abadia de Montserrat.
- Julià i Muné, Joan (2000c), *Pere Barnils: l'home, el lingüista i el mestre (1882–1933)*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Julià-Muné, Joan (2004a), *El llenguatge de la ràdio i de la TV*, Alzira, Bromera.
- Julià-Muné, Joan (2004b), *Francesc de B. Moll, lingüista aplicat*, in: Josep Massot (ed.), *Miscel·lània Joan Veny*, vol. 5, Barcelona, Publicacions de l'Abadia de Montserrat, 117–136.
- Julià-Muné, Joan (2005a), *La tasca gramatical d'Antoni M. Alcover*, *Llengua & Literatura* 16, 155–204.
- Julià-Muné, Joan (2005b, 2003), *The Development of Phonetic Sciences in the Catalan-speaking Countries during the Twentieth Century. A Chronological Approach to Phonetic and Phonological Studies of Catalan*, Lleida, Laboratori de Fonètica “Pere Barnils”.
- Julià-Muné, Joan (2005c), *Fonètica aplicada catalana. Dels fonaments a les aplicacions de les ciències fonètiques*, Barcelona, Ariel.
- Julià-Muné, Joan (2006), *Les contribucions lingüístiques al Primer Congrés Internacional de la Llengua Catalana: planificació de corpus versus planificació d'estatus*, *Estudis Baleàrics* 82/83, 183–206.
- Julià-Muné, Joan (2008), *Els primers “pensionats” per la llengua (1908–1912) i la lingüística catalana*, in: Josep Massot (ed.), *Homenatge a Joaquim Molas*, vol. 2, Barcelona, Publicacions de l'Abadia de Montserrat, 171–224.
- Julià-Muné, Joan (2010), *Sendas paralelas en torno a las ciencias fonéticas: Tomás Navarro (1884–1979) y Pere Barnils (1882–1933)*, in: José García-Velasco/José Manuel Sánchez Ron (edd.), *Actas del Congreso Internacional JAE 100. Centenario de la Junta de Ampliación de Estudios (Madrid, Febrero de 2008)*, Madrid, CSIC, 102–129.
- Julià-Muné, Joan (2012), *La fonamentació ortològica de Fabra i les aplicacions de Coromines*, in: Miquel Àngel Pradilla (ed.), *Fabra, encara*, Barcelona, Institut d'Estudis Catalans, 279–303.
- Labèrnia, Pere (3^a 1888–1892, 2^a 1864–1865, 1^a 1839–1840), *Diccionari de la llengua catalana ab la correspondència castellana y llatina*, 2 vol., Barcelona, Hereus de la V. Pla/Espasa Germans.
- Lacreu, Josep (6^a 2002, 1^a 1990), *Manual d'ús de l'estàndard oral*, València, Institut Interuniversitari de Filologia Valenciana, Universitat de València.
- Lamuela, Xavier/Murgades, Josep (1984), *Teoria de la llengua literària segons Fabra*, Barcelona, Quaderns Crema.
- Lloret, Maria-Rosa, et al. (edd.) (1997), *Anàlisi de la variació lingüística*, Barcelona, PPU.
- Marcet, Pere/Solà, Joan (1998), *Història de la lingüística catalana, 1775–1900. Repertori crític*, 2 vol., Vic/Girona, Eumo/Universitat de Girona.
- Martí i Castell, Joan (1990), *Gramàtica històrica. Problemes i mètodes*, València, Universitat de València.
- Martí i Castell, Joan/Pons, Lúdia/Solà, Joan (edd.) (1989), *Actes de les Terceres Jornades d'Estudi de la Llengua Normativa*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Mascaró, Joan (1978), *Catalan Phonology and the Phonological Cycle*, Bloomington, Indiana University, Linguistics Club. Catalan version: (1983), *La fonologia catalana i el cicle fonològic*, Bellaterra, Universitat Autònoma de Barcelona.
- Mascaró, Joan (1984), *Gabriel Ferrater i la tradició lingüística catalana*, *Els Marges* 31, 21–28.
- Mascaró, Joan (1986), *Morfologia*, Barcelona, Enciclopèdia Catalana.
- Mascaró, Joan (2006), *L'obra científica de Pompeu Fabra*, in: Pompeu Fabra (2005–2013), *Obres completes*, vol. 3, edd. Jordi Mir/Joan Solà, Barcelona/València/Palma, Proa/3i4/Moll, 13–54.
- Mascaró, Joan, et al. (edd.) (1984), *Estudis Gramaticals 1 (Working Papers in Linguistics/Universitat Autònoma de Barcelona)*, [Actes] del I Col·loqui Internacional de Lingüística Teòrica i Llengües

- Romàniques*, Bellaterra, Departament de Filologia Hispànica de la Universitat Autònoma de Barcelona.
- Massó i Torrents, Jaume/Casas-Carbó, Joaquim/Fabra, Pompeu (1891), *Campanya lingüística de "L'Avenç"*, *L'Avenç* (31.XII.1891), 375–389.
- Massot i Muntaner, Josep (1985), *Antoni M. Alcover i la llengua catalana*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Massot i Muntaner, Josep (1996), *Marià Aguiló, precursor d'Antoni M. Alcover*, in: Josep Massot (ed.), *Escriptors i erudits contemporanis*, Barcelona, Publicacions de l'Abadia de Montserrat, 11–17.
- Massot i Muntaner, Josep (2001), *Joan Coromines i Francesc de B. Moll*, in: Josep Massot (ed.), *Escriptors i erudits contemporanis. Segona sèrie*, Barcelona, Publicacions de l'Abadia de Montserrat, 185–203.
- Maugham, W. Somerset (1938), *The Summing Up*, London, Heinemann.
- Menéndez Pidal, Ramón (1902), *Cataluña bilingüe*, *El Imparcial* (Madrid, 15th Dec.).
- Menéndez Pidal, Ramón (1904), *Manual elemental de gramática histórica española*, Madrid, Revista de Filología Española.
- Meyer-Lübke, Wilhelm (1925), *Das Katalanische. Seine Stellung zum Spanischen und Provençalischen, sprachwissenschaftlich und historisch dargestellt*, Heidelberg, Winter.
- Milà i Fontanals, Manuel (1875), *Obras Completas. Estudios sobre historia, lengua y literatura de Cataluña*, vol. 3, Barcelona, Barcelonesa.
- Milà i Fontanals, Manuel (2^a1893, 1858), *Obras Completas. Opúsculos literarios, 2ª serie*, vol. 5, Barcelona, Barcelonesa.
- Miralles, Joan (1996), *Onomàstica i literatura*, Barcelona/Palma, Publicacions de l'Abadia de Montserrat/Universitat de les Illes Balears.
- Miralles, Joan (2001), *Entorn de la història de la llengua*, Barcelona/Palma, Publicacions de l'Abadia de Montserrat/Universitat de les Illes Balears.
- Miralles, Joan (2003), *Francesc de B. Moll: l'home dels mots*, Palma de Mallorca, Moll.
- Miralles, Joan (2005), *Francesc de B. Moll i la llengua literària*, Barcelona/Palma, Publicacions de l'Abadia de Montserrat/Universitat de les Illes Balears.
- Moll, Francesc de B. (1928–1931), *Suplement català al "Romanisches Etymologisches Wörterbuch"*, Barcelona, Balmes. Offprints of *Anuaris de l'Oficina Romànica de Lingüística i Literatura* 1928; 1929; 1930; 1931.
- Moll, Francesc de B. (1952), *Gramàtica històrica catalana*, Madrid, Gredos. Catalan version: (2006, 1991), *Gramàtica històrica catalana*, València, Universitat de València.
- Moll, Francesc de B. (1965), *Bernhard Schödel i els inicis de la lingüística catalana*, in: *Homenaje a Jaime Vicens Vives*, vol. 1, Barcelona, Universidad de Barcelona, Facultad de Filosofía y Letras, 95–110.
- Moll, Francesc de B. (1997, 1968, 1937), *Gramàtica catalana referida especialment a les Illes Balears*, Palma de Mallorca, Moll.
- Moll, Francesc de B. (1970), *Els meus primers trenta anys (1903–1934)*, Palma de Mallorca, Moll.
- Moll, Francesc de B. (1975), *Els altres quaranta anys (1935–1974)*, Palma de Mallorca, Moll.
- Moll, Francesc de B. (1980), *El parlar de Mallorca*, Palma de Mallorca, Moll.
- Moll, Francesc de B. (1981, 1962), *Un home de combat (Mn. Alcover)*, Palma de Mallorca, Moll.
- Moll, Francesc de B. (1982, 1959), *Els llinatges catalans*, Palma de Mallorca, Moll.
- Moll, Francesc de B. (1983a), *Aspectes marginals d'un home de combat (Mossèn Antoni M. Alcover)*, ed. Josep Massot, Barcelona, Curial/Publicacions de l'Abadia de Montserrat.
- Moll, Francesc de B. (ed.) (1983b), *Mossèn Antoni M. Alcover. Dietari de l'Excursió Filològica 1906*, Barcelona, Generalitat de Catalunya.
- Moll, Francesc de B. (2003–2006), *Obres completes*, ed. M. Pilar Perea, 3 vol., Palma de Mallorca, Moll.

- Moll, Francesc de B./Alcover, Antoni M. (1929–1932), *La flexió verbal en els dialectes catalans*, Anuari de l'Oficina Romànica de Lingüística i Literatura II–V.
- Moran, Josep (1994), *Treballs de lingüística històrica catalana*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Moran, Josep (1995), *Estudis d'onomàstica catalana*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Moran, Josep/Batlle, Mar/Rabella, Joan Anton (2002), *Topònims catalans. Etimologia i pronúncia*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Morel-Fatio, Alfred (1888), *Das Catalanische*, in: Gustav Gröber (ed.), *Grundriss der romanischen Philologie*, vol. 1, Strassburg, Trübner, 669–688.
- Moreu-Rey, Enric (1993), *Història dels nostres prenomes, cognoms i renoms*, Barcelona, Universitat de Barcelona.
- Murgades, Josep (1972), [Gabriel Ferrater 1922–1972] *El mestre*, Serra d'Or 14, 397–398.
- Murgades, Josep (1976), *Assaig de revisió del Noucentisme*, Els Marges 7, 35–53.
- Nadal, Josep M. (1992), *Llengua escrita i llengua nacional*, Barcelona, Quaderns Crema.
- Nadal, Josep M./Prats, Modest (1982), *Història de la llengua catalana*, vol. 1: *Dels orígens fins al segle XV*, Barcelona, Edicions 62.
- Nadal, Josep M./Prats, Modest (1996), *Història de la llengua catalana*, vol. 2: *El segle XV*, Barcelona, Edicions 62.
- Navarro Tomás, Tomás/Sanchis Guarner, Manuel (1934), *Análisis fonético del valenciano literario*, Revista de Filología Española 23, 113–141.
- Ninyoles, Rafael Lluís (¹1969), *Conflicte lingüístic valencià. Substitució lingüística i ideologies diglòssiques*, València, Eliseu Climent.
- Ninyoles, Rafael Lluís (¹1971), *Idioma i prejudici*, Palma de Mallorca, Moll.
- Ninyoles, Rafael Lluís (1976), *Bases per a una política lingüística democràtica a l'Estat espanyol*, València, 3i4.
- Ninyoles, Rafael Lluís (1989), *Estructura social i política lingüística*, Alzira, Bromera.
- Nomenclàtor = *Nomenclàtor oficial de toponímia major de Catalunya* (²2009, ¹2003), 3 vol., 2a ed., revisada i ampliada, Barcelona, Generalitat de Catalunya/Institut d'Estudis Catalans.
- Oliva, Salvador (1992), *La mètrica i el ritme de la prosa*, Barcelona, Quaderns Crema.
- Oliva, Salvador (2006), *Tractat d'elocució*, Barcelona, Empúries.
- Ors, Eugeni d' ("Xènius") (1982, ¹1906–1921), *Glosari*, ed. Josep Murgades, Barcelona, Edicions 62/"la Caixa".
- Palmada, Blanca (1994), *La fonologia del català. Els principis generals i la variació*, Bellaterra, Universitat Autònoma de Barcelona.
- Paloma, David/Rico, Albert (2000), *Diccionari de pronunciació en català*, Barcelona, Edicions 62.
- Par, Anfós (1923), *Sintaxi catalana segons los escrits en prosa de Bernat Metge (1398)*, Halle/Barcelona, Niemeyer/Foment de Pietat Catalana.
- Payrató, Lluís (1985), *La interferència lingüística. Comentaris i exemples català-castellà*, Barcelona, Curial/Publicacions de l'Abadia de Montserrat.
- Payrató, Lluís (¹1988), *Català col·loquial. Aspectes de l'ús corrent de la llengua catalana*, València, Universitat de València.
- Payrató, Lluís (1997), *De professió, lingüista. Panorama de la lingüística aplicada*, Barcelona, Empúries.
- Payrató, Lluís (ed.) (1998), *Oralment. Estudis de variació funcional*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Payrató, Lluís, et al. (edd.) (1996), *Corpus, corpora*, Barcelona, PPU.
- Perea, M. Pilar (1999a), *Complements a la flexió verbal dels dialectes catalans. Dotze quaderns de camp d'Antoni M. Alcover*, Barcelona, Publicacions de l'Abadia de Montserrat.

- Perea, M. Pilar (1999b), *Compleció i ordenació de "La flexió verbal en els dialectes catalans" d'A. M. Alcover i F. de B. Moll*, 2 vol., Barcelona, Institut d'Estudis Catalans.
- Perea, M. Pilar (ed.) (2003), *Francesc de B. Moll a l'inici del segle XXI*, Barcelona, Universitat de Barcelona.
- Perea, M. Pilar (2005), *Antoni M. Alcover dialectòleg, gramàtic, polemist*, Castelló de la Plana / Barcelona, Fundació Germà Colón/Publicacions de l'Abadia de Montserrat.
- Perea, M. Pilar (2006), *El centenari del Primer Congrés Internacional de la Llengua Catalana*, Barcelona, Biblioteca de Catalunya.
- Pérez Saldanya, Manuel (1988), *Els sistemes modals d'indicatiu i de subjuntiu*, Barcelona, Institut de Filologia Valenciana/Publicacions de l'Abadia de Montserrat.
- Pérez Saldanya, Manuel (1998), *Del llatí al català. Morfosintaxi verbal històrica*, València, Universitat de València.
- Petit i Aguilar, Joan (1998 [1796–1829]), *Gramàtica catalana*, ed. Jordi Ginebra, Barcelona, Institut d'Estudis Catalans.
- Pons, Arnau/Skrabec, Simona (edd.) (2008), *Carrers de frontera. Passatges de la cultura alemanya a la cultura catalana*, 2 vol., Barcelona, Institut Ramon Llull.
- Pradilla, Miquel Àngel (ed.) (1998), *El món dels sons*, Benicarló, Alambor.
- Prat de la Riba, Enric (1906a), *El Congrés d'avuy*, La Veu de Catalunya (13th October).
- Prat de la Riba, Enric (1978, 1906b), *La nacionalitat catalana*, Barcelona, Edicions 62/"la Caixa".
- Prieto, Pilar/Cabrè, Teresa (edd.) (2013), *L'entonació dels dialectes catalans*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Rafel i Fontanals, Joaquim (1981), *La lengua catalana fronteriza en el Bajo Aragón meridional: estudio fonológico*, Barcelona, Universidad de Barcelona.
- Rafel i Fontanals, Joaquim (ed.) (1996a), *Diccionari de freqüències*, vol. 1: *Llengua no literària*, Barcelona, Institut d'Estudis Catalans.
- Rafel i Fontanals, Joaquim (1996b), *El "Diccionari de l'Institut" i el "Diccionari Fabra"*, in: *Estudis de lingüística i filologia oferts a Antoni M. Badia i Margarit*, vol. 3, Barcelona, Publicacions de l'Abadia de Montserrat, 217–269.
- Rafel i Fontanals, Joaquim (ed.) (1998a), *Diccionari de freqüències*, vol. 2: *Llengua literària*, Barcelona, Institut d'Estudis Catalans.
- Rafel i Fontanals, Joaquim (ed.) (1998b), *Diccionari de freqüències*, vol. 3: *Dades globals*, Barcelona, Institut d'Estudis Catalans.
- Rasico, Philip D. (1982), *Estudis sobre la fonologia del català preliterari*, Barcelona, Curial/Publicacions de l'Abadia de Montserrat.
- Rasico, Philip D. (1993), *Estudis i documents de lingüística històrica catalana*, Barcelona, Curial.
- Rasico, Philip D. (2006), *El català antic*, Girona, CCG/UdG.
- Recasens, Daniel (1986), *Estudis de fonètica experimental del català oriental central*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Recasens, Daniel (1993), *Fonètica i fonologia*, Barcelona, Enciclopèdia Catalana.
- Recasens, Daniel (1996, 1991), *Fonètica descriptiva del català (Assaig de caracterització de la pronúncia del vocalisme i consonantisme del català al segle XX)*, Barcelona, Institut d'Estudis Catalans.
- Recasens, Daniel (2004), *A production account of sound changes affecting diphthongs and triphthongs in Romance*, *Diachronica* 21, 161–197.
- Recasens, Daniel (2011), *Velar and dental stop softening in Romance*, *Diachronica* 28, 186–224.
- Recasens, Daniel (2014), *Fonètica i fonologia experimentals del català*, Barcelona, Institut d'Estudis Catalans.
- Recasens, Daniel (2017), *Fonètica històrica del català*, Barcelona, Institut d'Estudis Catalans.
- Ribes, Salomé (2011), *L'obra lingüística d'Alfons Par*, Barcelona, Institut d'Estudis Catalans.

- Rico, Albert/Solà, Joan (1995), *Gramàtica i lexicografia catalanes: síntesi històrica*, València, Universitat de València.
- Rigau, Gemma (1975), *Linguistic Inquiry i la gramàtica generativa i transformacional*, Els Marges 4, 103–112; 5, 79–88.
- Rigau, Gemma (²1988, ¹1981), *Gramàtica del discurs*, Bellaterra, Universitat Autònoma de Barcelona.
- Robins, Robert H. (1967), *A Short History of Linguistics*, London, Longman.
- Roca-Pons, Josep (1971), *Introducció a l'estudi de la llengua catalana*, Barcelona, Vergara.
- Rossich, Albert (2006), *El model ortològic del català modern*, in: Antoni Ferrando/Miquel Nicolàs (edd.), *La configuració social de la norma lingüística a l'Europa llatina*, Alacant, Institut Interuniversitari de Filologia Valenciana, 125–153.
- Rossich, Albert/Rafanell, August (1998), *Oralitat, escriptura, ortologia*, in: *Actes de l'11è Col·loqui Internacional de Llengua i Literatura Catalanes*, Palma de Mallorca/Barcelona Universitat de les Illes Balears/Publicacions de l'Abadia de Montserrat, 263–298.
- Russell-Gebbett, Paul (ed.) (1965), *Mediaeval Catalan Linguistic Texts*, Oxford, The Dolphin Book.
- Sanchis Guarner, Manuel (1950), *Gramàtica valenciana*, València, Torre; reprinted (1993), ed. Antoni Ferrando, Barcelona, Alta Fulla.
- Sanchis Guarner, Manuel (⁴1972, ¹1933), *La llengua dels valencians*, València, L'Estel.
- Sanchis Guarner, Manuel (1980), *Aproximació a la història de la llengua catalana. Creixença i esplendor*, Barcelona, Salvat.
- Saroïhandy, Jean-Joseph (1904), *Das Catalanische*, in: Gustav Gröber (ed.), *Grundriss der romanischen Philologie*, vol. 1, Strassburg, Trübner, 841–877; partial reprint (1906), *Grammatik der katalanischen Sprache*, Strassburg, Trübner.
- Saroïhandy, Jean-Joseph (1908), *El català del Pirineu, a la ralla d'Aragó*, in: *Primer Congrés Internacional de la Llengua Catalana*, Barcelona, Joaquim Horta, 331–334.
- Schädel, Bernhard (1908), *Manual de fonètica catalana*, Cöthen, Schulze.
- Segarra, Mila (1985a), *Història de l'ortografia catalana*, Barcelona, Empúries.
- Segarra, Mila (1985b), *Història de la normativa catalana*, Barcelona, Enciclopèdia Catalana.
- Segarra, Mila (1998), *Pompeu Fabra. L'enginy al servei de la llengua*, Barcelona, Empúries.
- Solà, Joan (1972–1973), *Estudis de sintaxi catalana*, 2 vol., Barcelona, Edicions 62.
- Solà, Joan (1973), *Ortographe et grammaire catalane*, in: Antoni M. Badia i Margarit/Georges Straka (edd.), *La linguistique catalane. Colloque International organisé par le Centre de Philologie et Littératures Romanes de Strasbourg* (1968), Paris, Klincksieck, 81–100.
- Solà, Joan (1977a), *Del català incorrecte al català correcte*, Barcelona, Edicions 62.
- Solà, Joan (1977b), *A l'entorn de la llengua*, Barcelona, Laia.
- Solà, Joan (1987a), *L'obra de Pompeu Fabra*, Barcelona, Teide.
- Solà, Joan (1987b), *Qüestions controvertides de sintaxi catalana*, Barcelona, Edicions 62.
- Solà, Joan (1990), *Lingüística i normativa*, Barcelona, Empúries.
- Solà, Joan (1991), *Estudis d'història de la llengua catalana*, Barcelona, Empúries.
- Solà, Joan (²1994), *Sintaxi normativa: estat de la qüestió*, Barcelona, Empúries.
- Solà, Joan (ed.) (1999), *L'obra de Joan Coromines: cicle d'estudi i homenatge*, Sabadell, Fundació Caixa de Sabadell.
- Solà, Joan, et al. (edd.) (⁴2008, ¹2002), *Gramàtica del català contemporani*, 3 vol., Barcelona, Empúries.
- Ullastra, Josep (1980, [1743]), *Grammatica cathalana*, ed. Montserrat Anguera, Barcelona, Fundació Mediterrània.
- Vallverdú, Francesc (1968), *L'escriptor català i el problema de la llengua*, Barcelona, Edicions 62.
- Vallverdú, Francesc (1970), *Dues llengües, dues funcions?*, Barcelona, Edicions 62.
- Vallverdú, Francesc (1980), *Aproximació crítica a la sociolingüística catalana*, Barcelona, Edicions 62.
- Vallverdú, Francesc (1986), *Elocució i ortologia catalanes*, Barcelona, Junc.

- Vallverdú, Francesc (1990), *L'ús del català: un futur controvertit. Qüestions de normalització lingüística al llindar del segle XXI*, Barcelona, Edicions 62.
- Vallverdú, Francesc (2000), *El català estàndard i els mitjans audiovisuals*, Barcelona, TVC/Edicions 62.
- Veny, Joan (1973), *Dialectologie catalane*, in: Antoni M. Badia i Margarit/Georges Straka (edd.), *La linguistique catalane. Colloque International organisé par le Centre de Philologie et Littératures Romanes de Strasbourg* (1968), Paris, Klincksieck, 289–321.
- Veny, Joan (1978), *Estudis de geolingüística catalana*, Barcelona, Edicions 62.
- Veny, Joan (1986), *Introducció a la dialectologia catalana*, Barcelona, Enciclopèdia Catalana.
- Veny, Joan (¹²1998, ¹1978), *Els parlars catalans*, Barcelona/Palma de Mallorca, Dopesa/Moll.
- Veny, Joan (1999), *El mestratge d'Antoni M. Badia i Margarit*, *Revista de Llenguas y Literaturas Catalana*, Gallega y Vasca 6, 273–279.
- Veny, Joan (2002), *Treinta años de lingüística catalana*, in: Alberto Bernabé et al. (edd.), *Presente y futuro de la lingüística en España. Actas del II Congreso de la Sociedad Española de Lingüística*, Madrid, SEL, 291–324.
- Veny, Joan/Pons, Lúdia (1998), *Atles lingüístic del domini català. Etnotextos del català oriental*, Barcelona, Institut d'Estudis Catalans.
- Veny, Joan/Pons, Lúdia (2001–2017), *Atles lingüístic del domini català*, 9 vol., Barcelona, Institut d'Estudis Catalans.
- Veny, Joan/Pons, Lúdia (2007–), *Petit atles lingüístic del domini català*, Barcelona, Institut d'Estudis Catalans.
- Verdaguer, Pere (1974a), *Defensa del Rosselló català*, Barcelona, Curial.
- Verdaguer, Pere (1974b), *El català al Rosselló. Gal·licismes. Occitanismes. Rossellonismes*, Barcelona, Barcino.
- Verdaguer, Pere (1976), *Le catalan et le français comparés*, Barcelona, Barcino.
- Viana, Amadeu (1999), “Els Marges”: *un quart de segle*, *Serra d'Or* 479, 19–22.
- Viciano, [Rafael] Martí de (1877), *Alabanzas de las lenguas hebrea, griega, latina, castellana y valenciana*, València, Librería de Francisco Aguilar; facsimile edition (1979), València, Librerías París-Valencia.
- Wheeler, Max W. (1979), *Phonology of Catalan*, Oxford, Blackwell.
- Wheeler, Max W. (2005), *The Phonology of Catalan*, Oxford, Oxford University Press.
- Wheeler, Max W./Yates, Alan/Dols, Nicolau (1999), *Catalan. A Comprehensive Grammar*, London/New York, Routledge.

Appendix

A brief chronology of the Catalan-speaking regions

- 1881–1893 Pompeu Fabra and his friends Massó and Casas (*L'Avenç*) call for a new orthography of the language.
- 1901–1925 Flourishing Noucentisme under Catalan home rule
- 1906 The First Congrès Internacional de la Llengua Catalana is held in Barcelona. It is presided over by the Majorcan philologist Antoni M. Alcover. *La*

- nacionalitat catalana* by the President of the Diputació de Barcelona Enric Prat de la Riba is published.
- 1907 Prat de la Riba founds the Institut d'Estudis Catalans (IEC).
- 1912 Fabra's *Gramàtica de la lengua catalana* appears.
- 1913 The IEC approves the *Normes ortogràfiques*.
- 1914 Prat de la Riba presides over the Mancomunitat de Catalunya.
- 1917 The *Diccionari ortogràfic* by Fabra is published.
- 1918 Fabra's *Gramàtica catalana* appears. Its 7th edition (1933) will become the de facto official Catalan prescriptive grammar until 2016.
- 1919–1928 Fabra writes his newspaper columns under the title *Converses filològiques*.
- 1923–1930 Dictatorship of General Primo de Rivera
- 1925 The Mancomunitat is abolished and the IEC is suppressed.
- 1930 Alcover begins to publish his *Diccionari català-valencià-balear* (DCVB).
- 1931–1939 The Second Spanish Republic
- 1931 The Second Republic is proclaimed and King Alfonso XIII is forced into exile.
- 1932 The First *Estatut d'Autonomia de Catalunya* is approved and Francesc Macià becomes the President of the Generalitat de Catalunya. Fabra's *Diccionari general de la llengua catalana* (DGLC) is published in Barcelona.
- 1933 The Universitat Autònoma de Barcelona is founded and the historian Pere Bosch-Gimpera is appointed rector until 1939. Fabra heads its board of trustees.
- 1936–1939 The Spanish Civil War
- 1939 Franco's troops occupy Barcelona and the second dictatorship of the century is established. Fabra and Joan Coromines go into exile. Catalan is forbidden and the language remains absent from public life.
- 1939–1967 Post-war and exile phase
- 1941 Fabra's *Grammaire catalane* is published in Paris.
- 1948 Fabra dies in Prada de Conflent, in the Catalan-speaking area of southern France.
- 1950 The *Gramàtica valenciana* by Manuel Sanchis Guarner is published.
- 1954 The second edition of Fabra's *DGLC* appears.

- 1956 Coromines edits the posthumous *Gramàtica catalana* by Fabra.
- 1962 The *Obra Cultural Balear* is founded in Palma de Mallorca.
- 1964 *Llengua i cultura als Països Catalans* by Antoni M. Badia i Margarit is published.
- 1968–1977 The dawn of modernity
- 1968 Fabra's centenary and the Strasbourg conference on Catalan linguistics are the forerunners of a promising new era. Moll concludes the publication of the *DCVB*.
- 1969 Badia i Margarit publishes *La llengua dels barcelonins*.
- 1970 Amsterdam Conference on Catalan Linguistics.
- 1973 The Associació Internacional de Llengua i Literatura Catalanes (AILLC) is founded.
- 1975 General Franco dies and Juan Carlos I becomes King of Spain.
- 1977– The return of democracy: a revival of Catalan home rule
- 1977 The Generalitat de Catalunya is restored. Josep Tarradellas, current president of the Generalitat de Catalunya in exile, returns and is restored to his position by the King.
- 1978 Acció Cultural del País Valencià is founded by Joan Fuster and Eliseu Climent. Sanchis Guarner founds the Institut de Filologia Valenciana.
- 1979 The Second *Estatut d'Autonomia de Catalunya* is approved.
- 1980 Coromines starts publishing his *Diccionari etimològic i complementari de la llengua catalana*.
- 1983 The Language Normalization Act is passed in Catalonia. The Generalitat founds the Corporació Catalana de Ràdio i Televisió and its TV3 channel.
- 1989 The *Onomasticon Cataloniae* by Coromines begins to be published.

Language Description

Xavier Lamuela

3 Spelling

Abstract: Catalan spelling is a result of corpus planning carried out at the beginning of the 20th century in an atmosphere of political and linguistic vindication. This chapter first presents the criteria used for language codification and builds a framework for the analysis of alphabetic orthographies, including the different kinds of spellings used in them and the difficulties of reading and writing related to orthographic opacity and complexity. The chapter then goes on to discuss the spelling conventions of Catalan, taking into account possible coincidences with neighbouring languages, and classify them as morphological, diasystemic, historical, or etymological. Finally, the orthographic choices made for Catalan are interpreted with reference to codification criteria and language representations, and, in conclusion, it is observed that the corpus language planning activities were successful in achieving the aim of producing an orthography comparable to those of languages with a full range of uses characteristic of a modernized society.

Keywords: spelling, orthography, Catalan language, alphabetic, phonemic

1 Introduction

At the end of the 19th century Catalan had a flourishing literature but was excluded from formal public use (Anguera 1997). In 1895, Àngel Guimerà, a famous playwright, chose to deliver his first lecture as the new president of the cultural institution Ateneu Barcelonès in Catalan. This choice of language, which caused an enormous scandal (Anguera 1997, 216–217), was a sign of a new trend in the conception of the Catalan language as a central element of a large political and cultural movement in Catalonia (14.1 The Language Reform and the Work of Pompeu Fabra; Kremnitz 2018). The modernization of Catalan society during the 19th and 20th centuries led to the formation of a working class and an industrial bourgeoisie, which, in trying to consolidate its political influence through the use of the country's language, developed a cultural and linguistic policy as part of its political agenda. In 1914, Enric Prat de la Riba (1870–1917), a leader of the right-wing party Lliga Regionalista, succeeded in forming a single administration for the whole country of Catalonia out of the four pre-existing provinces, the Mancomunitat de Catalunya. A few years previously, in 1907, he had founded the Institut d'Estudis Catalans (IEC) as president of Barcelona's Provincial Council. It was the IEC which published the *Normes ortogràfiques* (1913) which, slightly modified in the *Diccionari ortogràfic* (1917), are essentially the current norms. They have twice been subject to marginal reforms: those implemented prior to the

<https://doi.org/10.1515/9783110450408-004>

publication of the new prescriptive *Diccionari de la llengua catalana* in 1995 (Institut d'Estudis Catalans 2007; see Institut d'Estudis Catalans 1997, 17–22) and those contained in the new *Ortografia catalana* (Institut d'Estudis Catalans 2017). The orthographic norms of 1913 were largely inspired by the work of Pompeu Fabra. Fabra (1948–1968) was also the author of the aforementioned *Diccionari ortogràfic* (1917), the prescriptive *Gramàtica catalana* (1918) and the *Diccionari general de la llengua catalana* (1932), the cornerstones of the Catalan language reform (see Fabra 2009 and 2005–2013; Ginebra/Solà 2007; Lamuela 1996; Lamuela/Murgades 1984; Segarra 1985a; 1985b; 1991).

2 Language representations and criteria for codification in their historical context

As Fabra's conception of what a "literary language" should be was decisive for the codification choices that were made, it is useful to outline its principal features. The notion of "literary language" as the language of written tradition is characteristic of the early 20th century and can, for example, be found in Saussure (1972, 267): "Par 'langue littéraire' nous entendons non seulement la langue de la littérature, mais, dans un sens plus général, toute espèce de langue cultivée, officielle ou non, au service de la communauté toute entière". It is also the key term in the Prague School's work on language cultivation: the development of the structural characteristics of languages that have the full range of uses distinctive of a given society, namely a modernized one, or what I have called "established languages". This was precisely Fabra's aim for Catalan (Cercle Linguistique de Prague 1929; Neustupný 1970; Scharnhorst/Ising 1976; Vachek/Dušková 1983; Lamuela/Murgades 1984, 35–40; Lamuela 1994, 106–109; Vila 2014, 58).

I established (Lamuela 1995) a list of possible criteria used in corpus language planning, added to later by Costa (2006, 2009). These include criteria that can be found in Fabra's work and will be used in section 6 to characterize Catalan orthography. "Literary languages", as conceived in Fabra's time, could entail some degree of rigidity, particularly resulting from the desire to avoid interferences from Spanish (the criterion of *autonomy*) and the relevance attributed to written tradition, including the medieval language (*historicity*), and the formal solutions characteristic of other European languages taken as models (*analogy with other languages*). This problem was partially counteracted by paying particular attention to the *acceptability* of the solutions proposed through the codification process and their implementation from a functional perspective, as is characteristic of the cultivation approach (*functionality*). The adoption of the Barcelona dialect as the basis for codification involved accepting its importance in terms of the sheer number of speakers and its position of prestige, though other dialects were also taken into account, and, in fact, a plural norm was

admitted. I refer to this way of dealing with dialects as the *diasystematicity* criterion in agreement with Bèc's notion (1972) of a diasystem as a group of dialects related to one another spatially or temporally. The principal methods for the tasks of corpus language planning were based on a knowledge of linguistics and guided by a concern for *distinctiveness* and *regularity*. We have then the following criteria: autonomy, historicity, analogy with other languages, acceptability, functionality, diasystematicity, distinctiveness, and regularity.

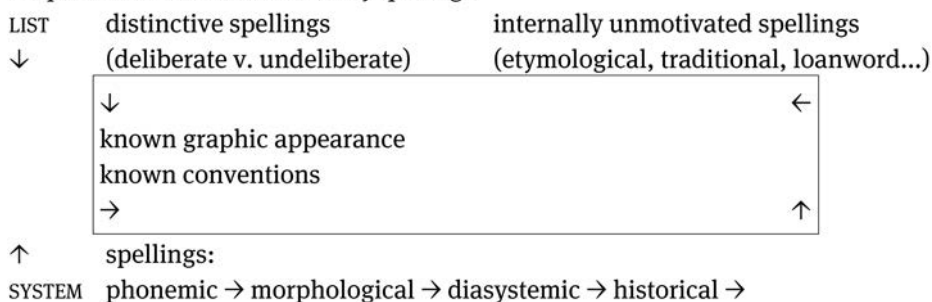
3 A framework for the analysis of alphabetic orthographies

It is generally believed that alphabetic orthographies should only use phonemic spellings, following what Venezky (1977, 37) has called “the highly suspicious principle” of one-to-one correspondence between phonemes and their written representations, or graphemes. Graphemes are the units of alphabetic writing systems, which may correspond to phonemes, be silent or have a particular function in spelling, such as determining the contextual value of another grapheme (Venezky 1999, 7). In reality, orthographies are always more complex than they are thought to be and even those that are considered to adhere closely to the phonemic principle rely greatly on the semantic interpretation of morphemes and whole words.

In two previous articles (Lamuela 1991; 2017), I started building a framework for the analysis and production of alphabetic orthographies. In the former, I produced a classification of the kinds of spellings that depart from the phonemic principle and discussed their effect on reading and writing difficulties. In the latter, I revised some aspects of the first one and developed the notions of opacity and complexity in orthographies, drawing mainly on Schmalz et al. (2015). The findings of this work constitute the basis for the following analysis of Catalan spelling.

3.1 Graphic conventions and kinds of spellings

Lafont (1971, 11–23, 31–38) presented a series of constraints that determine the way alphabetic orthographies work: 1. choice of graphemes, 2. inconsistencies in the use of graphemes, 3. language change, 4. paradigmatic morphology, 5. word-specific spellings, 6. loanwords, 7. word form, and 8. dialectal diversity (I have adapted the original terms). By reinterpreting and reorganizing these constraints, I produced the following diagram (Lamuela 1991, 73; 2017, 78), meant to explain the choice of graphic conventions and classify the different kinds of spellings:

Graphic conventions and kinds of spellings

The choice of writing conventions, comprising the choice of script, is influenced by cultural traditions. New written languages in Central and Western Europe adopted the Latin alphabet in the Middle Ages and its conventions had to be adapted to new sounds, as is the case with palatal sounds in the Romance languages, where, for example, [ɲ] is represented by <nh> in Portuguese and Occitan, and <gn> in French and Italian. The dependence of a new spelling on familiar conventions may lead to inconsistencies in their use (Lafont 1971, 17–18): thus, the adoption of <gn> for [ɲ] in French leads to ambiguity when it is used for [gn], as in *cognition* ‘cognition’ or *diagnostiquer* ‘to diagnose’. These are well-known facts and are very often discussed because of their relevant symbolic role in the look of orthographies; there is less awareness about the effects of known graphic appearance, although it also has an important symbolic function. People are reluctant to accept graphemes that alter the familiar aspect of written texts: for instance, it seems awkward to adopt <k> or use <ks> and <gz> instead of <x> in the orthography of a Romance language. One final issue related to the correspondence between phonemes and graphemes is that of Lafont’s word constraint (1971, 20–21); in most alphabetic orthographies, the representation of phonological changes in word boundaries is limited to a short list of function words, like the definite articles and clitic pronouns of French.

Several orthographies contain morphological spellings, representing morphemes rather than simple sequences of phonemes; these morphological spellings often ensure the stability of written forms (Lafont 1971, 19; Venezky 1999, 7–10; Catach ³2014, 17, 22–25, 203–260; Fayol/Jaffré 2014, 90–122). For example, the written form of lexical roots is constant in Portuguese, disregarding vowel reduction: *pedes* [ˈpɛðɐ] ‘(you) ask’, *pedir* [pɐˈðir] ‘to ask’; *podes* [ˈpɔðɐ] ‘(you) can’, *poder* [puˈðer] ‘to be able’. Schmalz et al. (2015, 1615) use the expression “morphological transparency”, as opposed to the “phonological transparency” that follows from the phonemic principle. Another kind of morphological spelling is used when written forms contain morphemes that are absent from or appear only sporadically in the spoken language, as is the case for certain verbal endings and the nominal plural suffix in French (Catach ³2014, 24–25). Catach (³2014, 17, 23, 201–260) uses the term “morphograms” to refer to morphological spellings.

Diasystemic spellings are common to different dialectal pronunciations. Lafont (1971, 31–38) devotes a number of pages to explaining what he calls the plurality constraint, according to which one particular spelling can be read following different dialectal pronunciations, as is the case for <z> or <c^{e/i}> in *zona* ‘zone’ or *cinco* ‘five’, read as [θ] in standard European Spanish and as [s] in most other varieties.

Because traditional orthographies are usually rather conservative, they often contain historical spellings that maintain the notation of old sounds. When the effects of vowel reduction are disregarded in writing, as in the case of Portuguese, it generally means that the current orthography reflects a previous linguistic stage, in agreement with the language change constraint defined by Lafont (1971, 18–19). Since language changes may concern only certain dialects, historical spellings are often also diasystemic.

Historical spellings, in the sense adopted here, are different from etymological ones. The former are derived from a previous stage of the same language, the latter from another language that may be its source but is viewed as separate. The use of <h> in French is etymological in words of Latin origin – such as *heure* ‘hour’ –, in which it has not been pronounced in any stage of the language. On the contrary, the use of <il(l)> for [j] corresponds to the pronunciation [ʎ], which was common until the 18th century in words like *paille* ‘straw’ or *travail* ‘work’. Etymological spellings are thus internally unmotivated, as are spellings resulting from loanwords being kept in their original form (Lafont 1971, 20) or those that derive from false etymologies – e.g. <d> and <g> in the French words *poids* ‘weight’ and *legs* ‘legacy’ – or various accidents in the history of a particular orthography, as is the case with the use of a final *x* in French words like *deux* ‘two’ or *chevaux* ‘horses’, deriving from a particular shape of the letter *z*. The presence of these kinds of spellings, as well as the deliberate use of certain distinctive marks, may result in particular graphic forms for specific words, or logograms (Catach 2014, 24; Lafont 1971, 19–20). A lot of French words are clearly distinguished in writing but are pronounced the same: *pois* [pwa] ‘pea’, *poix* [pitch], and *poids* ‘weight’. Accents may be used to differentiate words otherwise identical in spelling: e.g. Spanish *té* ‘tea’ and *te* ‘you (clitic)’.

Another aspect of spelling conventions is the use of a series of marks, like apostrophes, hyphens, accents, and diaereses. These are, in principle, a means of clarifying how a particular word is pronounced, as when accent marks are used to indicate stress or pitch. Apostrophes are often used to mark elisions or contractions, as in the English forms *don’t* or *couldn’t*. Hyphens help to distinguish either some clitics attached to another word – e.g. *levá-lo* in Portuguese or *portar-lo* ‘to bring it’ in Catalan – or the components of certain compound or prefixed words, as in English *pull-down* or *post-production*. So, it may be said in a general way that accents have a phonemic function and the use of apostrophes and hyphens contributes to morphological transparency.

The function of orthographies is characterized by a complex relation between general features and the characteristics of certain units, between what belongs to the

system and what has to be learnt from a list. It is well known from studies on the mechanisms of reading and writing that orthographies are used following a dual route, phonological and lexical, and that, when reading, adult native speakers use word recognition rather than grapheme-by-grapheme decoding (Cook 2004, 15–27; Fayol/Jaffré 2014, 55–89).

3.2 Opacity and complexity in orthographies

The expression “orthographic depth” is used to refer to the extent to which spellings depart from the phonemic principle (Katz/Frost 1992); the more they do, the deeper an orthography is considered to be. A deep orthography is expected to cause more difficulties in reading and writing than a shallower one. Schmalz et al. (2015) have distinguished two different aspects in orthographic depth: unpredictability, or opacity, and complexity. Here is an outline of the different cases of orthographic opacity and complexity (Lamuela 2017, 79–83).

Opacity and complexity in orthographies

1) Cases of opacity:

- Homographies: the use of one grapheme to represent more than one phoneme or the lack of representation of a phonemic unity → reading difficulties. Examples: the ambiguity of the digraph <ch> in French – *chœur* ‘choir’ /k/ and *choix* ‘choice’ /ʃ/ –, and the partial or complete absence of stress notation in many languages.
- Homophonies: the representation of one phoneme by more than one grapheme or the use of mute spellings → writing difficulties (spelling mistakes). Examples: the equal value of and <v> in Spanish – *acabar* ‘to end’ and *cavar* ‘to dig’ /b/ [β] – or <c> and <s> in French – *cent* ‘a hundred’ and *sens* ‘sense’ /s/.

Opacity requires the use of semantic information (lexical or morphological):

- Possibility of recognizing a word without context (logograms and morphograms). Example: the distinction between *cœur* ‘heart’ and *chœur* ‘choir’, both /'kœʁ/, in French.
- Need of context to recognize a word or a morpheme. Example: the Italian words *corso* /'kɔrso/ ‘Corsican’ and *corso* /'korso/ ‘course’ can only be distinguished in a given context, as in “un corso di corso” ‘a Corsican course’.

2) Cases of complexity in the formation and use of graphemes:

a) In the formation of graphemes:

- Use of diacritic signs: <å>, <ø>, <ç>, <č>.
- Use of digraphs, trigraphs...: <gn> for /ɲ/, <ll> for /ʎ/, <eu> for /ø/ or /œ/, <sch> for /ʃ/.

b) In the use of graphemes:

- Linear contextual values (reading rules): immediate context (Cook 2004, 13). Example: the distribution of the values /k/ and /s/ for <c> in the French words *cabane* ‘hut’ /k/ and *cire* ‘wax’ /s/.
- Non-linear contextual values (reading rules): non-immediate context (Cook 2004, 13). Example: the lengthening or diphthongization of English vowels in the context *Consonant* + <e> – *theme* ['θi:m], *lake* ['leɪk] – depend on the presence of <e>.
- Values (reading rules) depending on the grapheme position in words and phrases, taking into account prosodic phenomena. Example: reduction of unstressed vowels as in Portuguese – *pedir* [pə'ðir] ‘to ask’ and *poder* [pu'ðer] ‘to be able’.

4 General correspondences between sounds and graphemes in Catalan: spelling conventions

In this section, I shall present the values of Catalan graphemes and an outline of the use of accents, diaereses, apostrophes, and hyphens in Catalan orthography (Fabra 1956, §§ 1–23, § 35, § 39, §§ 62–65, § 82, §§ 154–163; Institut d'Estudis Catalans 1997; 2017; Wheeler/Yates/Dols 1999, 618–620).

Some observations on Catalan phonology and dialectal variation will be necessary to allow for a systematic explanation of Catalan spelling conventions (↗4 Phonology, Phonetics, Intonation; ↗8.1 Dialects; Recasens 1991; Veny/Massanell 2015). The basis for my analysis will be the system of the prescriptive variety based on the Barcelona dialect. The inventory of phonemes, though not their distribution, is the same for most dialects. The important exceptions here are the use of /dʒ/ instead of /ʒ/ in most Valencian varieties and Alguerese, the lack of voiced affricates and fricatives in Central Valencian and some peripheral varieties spoken in Aragon, the presence of /v/ in Balearic, Alguerese, and some Valencian varieties, and the use of /ə/ with phonemic value in Balearic.

In the organization of Catalan orthography, /k^w/ and /g^w/ are treated as phonemes, as will be seen when dealing with the use of accent marks. Depending on the phonological theory adopted, /ŋ/ may be considered a phoneme on the basis of oppositions such as that between *sant* ['san] ‘saint’ and *sang* ['san] ‘blood’. It can be argued that /j/ and /w/ are autonomous phonemes with the realizations [j] / [i] and [w] / [u], respectively; considering them as phonemes is also a practical solution for explaining Catalan orthography. Breaking with the usual practice, I mark the fact that the trill is long, /r:/, when syllable initial.

Here is an annotated list of Catalan orthographic conventions, ordered according to the manner of articulation of consonants and the openness of vowels. As usual, graphemes are written between angle brackets, <p>. Where Catalan conventions

coincide with those generally used in the Romance languages, I omit this from my comments. By “neighbouring languages” I mean Occitan, French, and the Romance languages spoken in the Iberian Peninsula; if Basque is included, I mention it explicitly.

Consonants

Stops

/p/ – <p> <-b>	pare ‘father’, tap ‘stopper’, àrab ‘Arab, Arabic’
Ø – <-p> <-b>	camp ‘field’, rumb ‘route’
/b/ – <-v>	bo ‘good’, vi ‘wine’, roba [β] ‘clothes’, cavall [β] ‘horse’
[b:t] – <bl>	amable ‘kind’
/t/ – <t> <-d>	terra ‘earth’, pot ‘pot’, fred ‘cold’
Ø – <-t> <-d>	alt ‘high’, dent ‘tooth’, horts ‘gardens’, herald ‘herald’, profund ‘deep’, covards ‘cowards’
/d/ – <d>	dit ‘finger’, roda [ð] ‘wheel’
/k/ – <c> <-g> <qu ^{e/i} >	car ‘expensive’, quiet ‘still (adjective)’, sac ‘sack’, pròfug ‘fugitive’, càstig ‘punishment’
/g/ – <g> <gu ^{e/i} >	gat ‘cat’, guiar ‘to guide’, pagar ‘to pay’ [ɣ], ceguesa [ɣ] ‘blindness’
[g:t] – <gl>	segle ‘century’
/k ^w / – <qu> <qü ^{e/i} >	quatre ‘four’, conseqüència ‘consequence’
/g ^w / – <gu> <gü ^{e/i} >	guardar ‘to save’, pingüí ‘penguin’, aigua ‘water’ [ɣ ^w], aigües ‘waters’ [ɣ ^w]

The value /k/ of <c> alternates with /s/, found before front vowels, <c^{e/i}>, as is the case generally in the neighbouring languages. <qu> and <gu> are used to represent /k/ and /g/, respectively, before front vowels, as in the other Western Romance languages. <qu> is used for /k^w/, as in Portuguese and unlike in Spanish.

There is spirantization of the voiced stops after a continuant sound: /b/, /d/, /g/, and /g^w/, into [β], [ð], [ɣ], and [ɣ^w], respectively.

Obstruents are always voiceless in word final position; final stops are represented according to etymology. The etymological spelling has been adopted in Occitan from Catalan (Alibèrt 1935).

In some varieties, including those spoken in Catalonia, word final stops are silent after a nasal, /t/ also after /t/; <nc> and <ng> are pronounced [ŋ] – see below. /t/ between /r/ and /s/ is usually omitted.

/v/ is a phoneme different from /b/ in the areas already mentioned above.

Fricatives

/f/ – <f>	farina ‘flour’
(/v/ – <v>)	vi ‘wine’, cavall ‘horse’
/s/ – <s> <-ss-> <c ^{e/i} > <ç>	sac ‘sack’, pas ‘step’, pensar ‘to think’, passar ‘to pass’, cel ‘sky’, caçar ‘to hunt’, braç ‘arm’
/ks/ – <x>	fixa ‘fixed f.’
/z/ – <-s-> <z>	casa ‘house’, zero ‘zero’, colze ‘elbow’, ozó ‘ozone’
/gz/ – <x>	examen ‘exam’
/ʃ/ – <x-> <-ix(-)>	xoc ‘crash’, panxa ‘belly’, rauxa ‘outburst’, mixa ‘puss’, caixa ‘box’, gruix ‘thickness’
/ʒ/ – <j> <g ^{e/i} >	roja ‘red f.’, gel ‘ice’, objecte ‘object’

The use of the alternations <-ss->, intervocalic, / <s>, elsewhere, and <c^{e/i}> / <ç>, for /s/, and <-s->, intervocalic, / <z>, everywhere, are found, albeit with some differences in distribution, in French, Occitan, and Portuguese. The spelling for /ʒ/ – or /dʒ/ – is also <j> / <g^{e/i}> in the same languages. The use of <(i)x> for /ʃ/ is historically characteristic of the languages spoken in the Iberian Peninsula, including Basque, and adjacent Gascon.

Affricates

/ts/ – <t>	pots ‘pots’, potser [t:s] ‘maybe’
/dz/ [d:z] – <tz>	dotze ‘twelve’
/tʃ/ – <tx> <-ig>	despatx ‘office’, despatxa ([t:ʃ]) ‘(he...) dispatches’, roig (roja f.) ‘red’, mig (mitja f.) ‘half a’
/dʒ/ [d:ʒ] – <tj> <tg ^{e/i} >	corretja ‘strap’, viatge ‘trip’

<t> is used to mark affrication even with voiced sounds. Final <ig> for /tʃ/ is used when alternating with <j> / <g^{e/i}> or <tj> / <tg^{e/i}>. Occitan shares with Catalan the use of <tz>, <tj> / <tg^{e/i}>, and final <g>, without <i>, for /tʃ/ – e.g. *puèg*, Catalan *puig* ‘peak’.

Nasals

/m/ – <m>	mico ‘monkey’, rem ‘oar’
/n/ – <n>	dona ‘woman’
/ɲ/ – <ny>	canya ‘reed’
/ŋ/ – <-nc> <-ng>	blanc ‘white’, sang ‘blood’

The spelling <ny> for /ɲ/ is specific to Catalan in its geographical area but present in other languages.

Laterals

/t/ – <l>	tela ‘cloth’
/t:/ – <-l-l->	col·legi ([t]) ‘school’, atles ([dɫ]) ‘atlas’ (ametla [t:] / ametlla /ɫ:/ ‘almond’)
/ʎ/ – <ll>	llop ‘wolf’, pell ‘skin, fur’
/ɫ:/ – <tll>	ametlla (/ ametla /t:/) ‘almond’

The use of <ll> for /ʎ/ is characteristic of the languages spoken in the Iberian Peninsula, except for Portuguese. The spelling /tl/ for /t:/ has its basis in phonetic evolution and is shared with Occitan. It is kept in most Catalan dialects, in some places representing a short /t/; in Catalonia, except for in the South, it has been adapted to represent a new evolution of double *l* towards palatalization: /ɫ:/, represented by <tll>. The use of <ll> for /ʎ/ poses a problem of “inconsistency in the use of graphemes” (Lafont 1971, 18) when it comes to representing /t:/; then, the solution <-l-l-> was devised.

Rhotics

/r/ – <r>	cara ‘face’, prat ‘meadow’
/r/ [r] / [r] – <r>	mar ‘sea’, porta ‘door’
/r:/ – <-r-> <-rr->	roc ‘stone’, enreixat ‘grating’, torre ‘tower’
Ø <-r>	por ‘fear’, cantar ‘to sing’

The languages spoken in the Iberian Peninsula, including Basque, as well as Southern Occitan share the same spelling conventions for what is, or was formerly, a similar phonological situation. Final <r> is generally silent in most Catalan and Occitan dialects.

Semi-vowels

/j/ [j] / [j] – <i>	iode ‘iodine’, noia ‘girl’, mai ‘never’
/w/ [w] / [w] – <u>	diuen ‘they say’, cau ‘den’

The use of <y> for /j/ is avoided in Catalan as in Portuguese, Galician, Basque, and Occitan, and unlike in Spanish, Asturian, Aragonese, and French.

Silent

Ø – <h>	home ‘man’

<h> is used on etymological grounds in the neighbouring languages, except for Occitan and, partially, Aragonese.

*Vowels**Close*

/i/ – <i> <í> <i>	pi ‘pine’, pingüí ‘penguin’, veïna ‘neighbour f.’
/i/ – <i> <ï>	mirar ‘to look’, canti ‘sing (polite)’, veïnat ‘neighbourhood’
/u/ – <u> <ú> <ü>	mut ‘mute’, búfal ‘buffalo’, diürn ‘diurnal’
/u/ – <u> <ü> <o>	durar ‘to last’, diürètic ‘diuretic’, ferro ‘iron’, gotera ‘leak’

When, according to the accent rules explained below, <i> and <u> carry an accent, it is always acute.

Apart from its use in the sequences <qü^e/i> and <gü^e/i> to ensure the pronunciations [k^w] and [g^w], a diaeresis is used in Catalan orthography to mark <i> and <u> as vowels (Institut d’Estudis Catalans 2017, 101–103; Wheeler/Yates/Dols 1999, 613–614): *veïna* [bə.ˈi.nə] ‘neighbour f.’ ~ *reïna* [ˈrɛi̯.nə] ‘queen’. In some words, where the pronunciation in two syllables is considered predictable, the use of the diaeresis is avoided. The same convention is used in Occitan, whereas an accent fulfils this function, limited to stressed vowels, in Spanish and Portuguese.

Vowel reduction is not reflected in writing; so, an unstressed <o> is read as [u] in most Eastern dialects. A similar vowel reduction exists in Occitan and European Portuguese.

Close-mid and open-mid

/e/ – <e> <é>	teu ‘yours’, excés ‘excess’
/e/ – <e>	teatre [teˈatrə] ‘theatre’, teatral [teˈatral] ‘theatrical’
/o/ – <o> <ó>	gota ‘drop’, balcó ‘balcony’
/o/ – <o>	plourà [pɫuɾa] ‘it will rain’
/ɛ/ – <e> <è>	peu ‘foot’, pagès ‘countryman’
/ɔ/ – <o> <ò>	cosa ‘thing’, això ‘this, that’

<e>, when stressed without an accent mark, may represent /e/ and /ɛ/; <o>, in the same conditions, may represent /o/ and /ɔ/. <e> is also the notation for /ə/ in the

varieties that have this phoneme. When <e> and <o> bear an accent, it is acute – <é>, <ó> – for /'e/ and /'o/, and grave – <è>, <ò> – for /ε/ – also for /'ə/ – and /'ɔ/. This use of acute and grave accents coincides with the French one for <e> and with that of Occitan and Italian for both <e> and <o>, although in Occitan an ancient /o/ has become /u/. When Western dialects have /'e/ where the Eastern ones have /'ε/, both shapes of accent are allowed in most words: *anglés* ‘English’ (Western use) – *anglès* (Eastern use).

[e] and [o] are the result of vowel reduction in Western dialects. Unstressed [e] and [o] are present in some contexts in Eastern dialects. Majorcan has unstressed [o] and has re-established unstressed /e/ in some cases: *pecar* [pe'ka] ‘to sin’.

Mid

[ə] – <e> <a>	porta ‘door’, país ‘country’, home ‘man’, pesar ‘to weigh’

[ə] is the result of vowel reduction of unstressed /a/, /ε/, and /e/ in most Eastern dialects. /ə/ is a phoneme that is found in stressed position in most Balearic varieties. European Portuguese has [ə], written <e> in unstressed position.

Open

/'a/ – <a> <à>	camp ‘field’, àrab ‘Arab, Arabic’

<a> takes the grave accent, <à>, when the rules require one.

Most orthographies of languages with variable stress do not mark stressed syllables. In Spanish, Portuguese, Occitan and Catalan, they are systematically indicated through complementary rules; stress is only marked when it falls on an unexpected syllable. So, in Catalan, the following rules apply (Institut d’Estudis Catalans 2017, 87–96; Wheeler/Yates/Dols 1999, 614–616):

- a) In words ending in a vowel, a vowel plus s, or in *-en* or *-in*, the stress is not marked when it falls on the next to last syllable – *Pasqua* ‘Easter’, *aigua* ‘water’, *casa* ‘house’, *cases* ‘houses’, *canten* ‘(they) sing (indicative)’, *cantin* ‘(they) sing (subjunctive)’ – and is marked if it falls on the last syllable – *camí* ‘path’, *cabàs* ‘basket’, *encén* ‘(he...) lights’.
- b) In words with an ending different from those in (a), the stress is not marked when it falls on the last syllable – *fidel* ‘faithful’, *creieu* ‘(you) believe pl.’, *segon* ‘second’ – and is marked if it falls on the next to last syllable – *àngel* ‘angel’, *crèieu* ‘(you) believed pl.’, *cànon* ‘canon (norm)’.
- c) The stress is always marked when it falls on the second to last syllable: *ràpida* ‘rapid f.’, *història* ‘history’, *vàlua* ‘worth’.

These rules work on the two assumptions that sequences of <i> or <u> followed by a mid or open vowel are always disyllabic – *història* [is.ˈtɔ.ri.ə], *vàlua* [ˈba.ɫu.ə] –, which is not the case for most speakers, and that, on the contrary, <u> in <qu> and <gu> does not represent a vowel – *aigua* [ˈaj.ɣ̞ːwə], *Pasqua* [ˈpas.kwə].

It has to be said that, although such rules guarantee a phonemic representation of stress, they are rarely mastered by the users because of the need to learn them explicitly, since it is impossible to internalize them from practice.

Accents are also used with a distinctive function (“diacritic accents”) to yield particular cases of logograms, as in *mà* ‘hand’, different from *ma* ‘my f.’. The most recent orthographic reform has reduced the list of such logograms to fifteen, most of them with their plurals where possible (Institut d’Estudis Catalans 2017, 96–99).

Apostrophes are used to indicate vowel elision in definite and personal articles, clitic pronouns, and the preposition *de* ‘of’ (Institut d’Estudis Catalans 2017, 105–109; Wheeler/Yates/Dols 1999, 43–46, 167–170, 204–205): *l’home* ‘the man’ (*el* ‘the m.sg.’ → *l’*), *l’estació* ‘the station’ (*la* ‘the f.sg.’ → *l’*), *n’Àngela*, *n’Enric* (Balearic Catalan: *na f.* → *n’*, *en m.* → *n’*), *l’escolta* ‘(he...) listens to it / him / her’ (*la* ‘3rd f.sg.’ / *el* ‘3rd m.sg.’ → *l’*), *dona-me’l* ‘give it to me’, *manual d’història* ‘history textbook’ (*de* → *d’*).

Hyphens are found between verbal forms and enclitic pronouns – if the use of apostrophe does not apply – as well as in numerals and compounds (Institut d’Estudis Catalans 2017, 111–127; Wheeler/Yates/Dols 1999, 618–620): *portar-m’ho* ‘to bring it to me’, *trenta-dos* ‘thirty-two’, *sud-americà* ‘South American’, *abans-d’ahir* ‘the day before yesterday’, *pit-roig* ‘robin’, *penya-segat* ‘cliff’, *para-xocs* ‘bumper’. The hyphen in the last three examples is meant to clarify the pronunciation of <r>, <s>, and <x> as /r:/, /s/, and /f/, respectively.

5 Kinds of spellings in Catalan: morphological, diasystemic, historical, etymological, and distinctive

In section 3.1 we saw the different kinds of spellings possible in alphabetic writing systems. The following table presents Catalan spelling choices marked according to their type: morphological, diasystemic – which are also historical –, or etymological. My comments on these choices are below the table.

Kinds of spellings in Catalan

Spellings	Morpho-logical	Diasys-temic	Etymo-logical
01 <p> /p/ àrab (aràbic) ‘Arab, Arabic’ ~ rep (rebre) ‘(he...) receives (to receive)’	–	–	+
02 <bt> /pt/ dissabte [pt] ‘Saturday’ ~ apte ‘fit’	–	–	+
03 <rp> /rp/ serp [rp] ‘snake’ – serpeta [rp] ‘little snake’	+	–	+
04 <rb> /rp/ verb [rp] ‘verb’ – verbal [rβ] ‘verbal’	+	–	+
05 <mp> /m/ camp [m] ‘field’ – acampar [mp] ‘to camp’	+	+	+
06 <pn-> /n/ pneumàtic [n] ‘pneumatic, tyre’	–	–	+
07 <pt-> /t/ pterosaure [t] ‘pterosaur’	–	–	+
08 <ps-> /s/ / /ps/ psicòleg [s] / [ps] ‘psychologist’	–	–	+
09 <mb> /m/ tomb [m] ‘turn’ – tombar [mb] ‘to turn’	+	+	+
10 <v> /b/ vi [b] ‘wine’, cantava [β] (he...) ‘sang’	–	+	–
11 <v> /b/ blava [β] – blau [u] ‘blue f. – m.’	+	+	–
12 <t> <d> /t/ fred (freda) ‘cold m. (f.)’ ~ fat (fada) ‘tasteless m. (f.)’	–	–	+
13 <tm> /dm/ ritme [dm] ‘rhythm’ ~ administrar ‘to administer’	–	–	+
14 <rt> /rt/ fort [rt] – forta [rt] ‘strong m. – f.’	+	–	+
15 <rd> /rt/ verd [rt] – verda [rð] ‘green m. – f.’	+	–	+
16 <nt> /n/ font [n] ‘fountain’ – fonteta [nt] ‘little fountain’	+	+	+
17 <nd> /n/ rotund [n] – rotunda [nd] ‘emphatic m. – f.’	+	+	+
18 <lt> /t/ alt [t] – alta [t̺t] ‘high m. – f.’	+	+	+
19 <ld> /t/ herald [t] ‘herald’ – heràldic [t̺d] ‘heraldic’	+	+	+
20 <c> <g> /k/ mag (maga) ‘magician m. (f.)’ ~ amic (amiga) ‘friend m. (f.)’	–	–	+
21 <cm> /gm/ dracma ‘drachma’ [gm] ~ magma ‘magma’	–	–	+
22 <rc> /rk/ arc [rk] ‘arch’ – arcada [rk] ‘arcade’	+	–	+
23 <rg> /rk/ alberg ‘hostel’ [rk] – albergar [ry] ‘to shelter’	+	–	+
24 <nc> /ŋ/ blanc [ŋ] – blanca [ŋk] ‘white m. – f.’	+	+	+
25 <ng> /ŋ/ fang [ŋ] ‘mud’ – fangós [ŋg] ‘muddy’	+	+	+
26 <-s-> /s/ asimètric [s] ‘asymmetric’ ~ bassa ‘pond’	+	–	+
27 <c ^{e/i} > <ç> /s/ cel ‘sky’, caçar [s] ‘to hunt’ ~ sopa ‘soup’	–	–	+
28 <x> /gz/ /ks/ examen [gz] ‘exam’, fix [ks] ‘fixed’ ~ rocs ‘stones’	–	–	+
29 <z> /z/ amazona [z] ‘Amazon’ ~ casa ‘house’	–	–	+

Spellings	Morpho-logical	Diasys-temic	Etymo-logical
30 <j ^e > /3/ objecte ‘object’ [3] ~ gel ‘ice’	–	–	+
31 <ig> /tʃ/ roig [tʃ] – roja [3] ‘red m. – f.’	+	–	–
32 <mpt> /nt/ compte [nt] ([mt]) ‘account’ ~ conte ‘tale’, manta ‘blanket’	–	–	+
33 <mf> /nf/ àmfora [mf] ‘amphora’ ~ informe ‘report’	–	–	+
34 <l·l> /t/ pel·lícula [t] ‘film’ ~ sola ‘alone f.’	–	–	+
35 <l·l> /t/ (in some words) cel·la ‘cell’	–	+	+
36 <-r> Ø dur Ø – dura [r] ‘hard m. – f.’	+	+	+
37 <-r> /r/ contrarestar [r] ‘to counteract’ ~ carro ‘cart’	+	–	+
38 <h> Ø home ‘man’	–	–	+
39 <o> /u/ ferro ‘iron’, gotera [u] ‘leak’ – gota [‘o] ‘drop’	+	+	–
40 <e> /a/ [ə] home ‘man’, sec [‘ε] ‘dry’ – secor [ə] ‘dryness’	+	+	–
41 <a> /a/ [ə] porta ‘door’, blanc [‘a] ‘white’ – blancor [ə] ‘whiteness’	+	+	–

The application of an etymological principle is apparent in the solution adopted for final voiceless stops (points 1, 3–5, 9, 12, 14–20, 22–25). When they follow a consonant, the etymological spelling yields an effect of morphological transparency (3–5, 9, 14–19, 22–25). In some dialects, namely in Barcelona, the final stop is silent depending on the preceding consonant (5, 9, 16–19, 24–25) but has diasystemic correspondences. The choice of writing certain clusters is also etymological: initial – e.g. <pn>, <pt>, <ps> (6–8), from which only the third is pronounced as a cluster, [ps], by some speakers –, or internal – e.g. <bt>, <tm>, <cm>, <mpt>, <mf> (2, 13, 21, 32–33). A particular case is that of *compte* ‘account’, distinguished from *conte* ‘tale’, which has the same etymon, and *comte* ‘earl, count’; the distinction, imitated from French, was initially only orthographic, but the pronunciation [‘kومتə] for *compte* and *comte* is now common. Other etymological choices are <x> for /ks/ and /gz/ (28), intervocalic <z> for /z/ (29), <j^e> instead of <g^e> for /3/ (30), silent <h> (38), <l·l> for /t/ (34) – although for some words, where it represents /t:/, particularly in Balearic, it can be considered diasystemic or even phonemic (35). The distribution of <c^e/i> and <ç> for /s/ was chosen on etymological, not historical or diasystemic grounds (27). The use of a simple intervocalic <s> for /s/ (26) or <r> for /r:/ (37) on morpheme borders corresponds to etymological criteria but has an effect of morphological transparency.

Final <ig> for /tʃ/ alternating with /3/ or /d3/ (31) is a traditional spelling but works as morphological by using a grapheme related to those found between vowels.

<v> for /b/ (10) is written according to the ancient texts and the pronunciation of the dialects that distinguish /v/ from /b/ and not according to etymology, as can be

seen in some cases: *buit* ‘empty’ (from VOCITU), *canviar* ‘to change’ (from CAMBIARE). The alternation of final <u> [u] with intervocalic <v> constitutes a morphological spelling (11).

The spellings that avoid representing vowel reduction (39, 40, 41) are morphological, historical, and diasystemic, but some choices may be influenced by etymology, like <e> in *sencer* ‘complete’ or *lleuger* ‘light (adjective)’, since the common pronunciation in Western dialects is [san’sɛ(r)] or [lɐu’(d)ʒɛ(r)].

Final <r>, silent in most dialects, is written according to morphological transparency, written tradition, diasystem, and etymology (36).

As for complementary marks, accents indicating stress and diaereses have a phonemic function, diacritic accents a distinctive one. The use of apostrophes and hyphens normally highlights morphological transparency. Apostrophes, on the other hand, contribute to the representation of phonemic variation when the word constraint is not fulfilled; they have then a phonemic function. Hyphens also have a phonemic function when they are used to clarify the pronunciation of <r>, <s>, and <x> as /r:/, /s/, or /ʃ/, as in the examples at the end of section 4.

6 Language representations and orthographic choices

As mentioned in section 2, the criteria that guided Fabra’s codification of Catalan, which reflect a particular conception of language in society, are recognizable in the organization of the orthography.

The choice of graphic conventions corresponds to Catalan tradition, according to principles of autonomy and historicity. It makes them acceptable from a symbolic point of view as well as from a functional one, since, in the case of Catalan, and more precisely of Catalonia, they were still in use at the time of the language reform. Apart from <ny> and the innovation <l·l>, motivated by the need to avoid confusion with <ll> used for /ʎ/, all the conventions are used in neighbouring languages. This fact and the desire to keep the known graphic appearance of written words are clearly related to the criterion of analogy with other languages, which is also responsible for etymological choices that usually increase the number of homophonies and spelling difficulties but make the orthography more acceptable to some speakers on symbolic grounds. The fact that there are few homographies reflects, on the contrary, a concern for systematicity and regularity in codification work.

There are no explicitly morphological choices in Catalan orthography apart from the cases where a prefix and a lexical root are kept distinct: *contrarestar* ‘to counteract’, where <-r> /r:/, or *contrasenya* ‘password’, where <-s> /s/. But a lot of diasystemic and etymological spellings yield morphological transparency, as is the case, respectively, for the alternations <r> Ø / <r> /r/ – *dur* ‘hard m.’ / *dura* ‘hard f.’ – and

<d> /t/ / <d> /d/ [ð] – *verd* ‘green m.’ / *verda* ‘green f.’. In one case, a traditional spelling is also morphological: <ig> /tʃ/ / <j> /ʒ/ – *roig* ‘red m.’ / *roja* ‘red f.’.

Diasystemic and historical spellings, both characteristic of Catalan orthography, are a guarantee of common practice and the cultural use of written tradition, which are fundamental for the success of any language reform from a symbolic as well as from a functional point of view.

A concern for regularity and distinctiveness, related to a “linguist’s mentality”, may explain the widespread use of apostrophes, hyphens, diaereses and accents, intended to clarify the pronunciation of every word, but which results in considerable spelling difficulties.

7 Conclusions

The result of a language reform linked to a process of linguistic vindication, Catalan spelling is comparable to those of established languages, both in its positive and negative characteristics. It has the general appearance of the orthographies of the other Romance languages and satisfactorily represents the language in its different forms, both historical and dialectal. As the model of established languages was successfully adopted in Catalan orthography, it is capable of fulfilling the different cultural and social functions expected from an orthography in the present world. From the point of view of transparency and practical efficiency, Catalan orthography makes the language fairly easy to read correctly, but creates an excessive amount of writing difficulties. These are partly due to diasystemic spellings, which is a reasonable price to pay to ensure common practice of the language. But etymological and unnecessary distinctive spellings also play a role in orthographic difficulties we would be better off without.

8 Bibliography

- Alibèrt, Loís (1935), *Gramatica occitana según los parlars lengadocians*, Toulouse, Societat d’Estudis Occitans.
- Anguera, Pere (1997), *El català al segle XIX. De llengua del poble a llengua nacional*, Barcelona, Empúries.
- Bèc, Pèire (1972), *Per una dinamica novèla de la lenga de referència: dialectalitat de basa e diasistèma occitan*, Annales de l’Institut d’Études Occitanes 6, 4th series, 39–61.
- Catach, Nina (2014, 1980), *L’orthographe française: traité théorique et pratique avec des travaux d’application et leurs corrigés*, avec la collaboration de Claude Gruaz/Daniel Duprez, Paris, Colin.
- Cercle Linguistique de Prague (1929), *Thèses*, in: *Mélanges linguistiques dédiés au Premier Congrès des Philologues Slaves*, Prague, Jednota Československých Matematiků a Fyziků, 7–29.
- Cook, Vivian (2004), *The English Writing System*, London, Arnold.

- Costa, Joan (2006), *Criteria for Linguistic Codification and Completion*, in: Vittorio Dell'Aquila/Gabriele Iannàccaro/Matthias Stuflesser (edd.), *Alpes Europa 2. Soziolinguistica y language planning*, Trento, Regione Autonoma Trentino-Alto Adige/Istitut Cultural Ladin/Centre d'Études Linguistiques pour l'Europe, 36–53.
- Costa, Joan (2009), *La norma sintàctica del català segons Pompeu Fabra*, Munich, Peniöpe.
- Fabra, Pompeu (1917), *Diccionari ortogràfic*, Barcelona, Institut d'Estudis Catalans.
- Fabra, Pompeu (1918), *Gramàtica catalana*, Barcelona, Institut d'Estudis Catalans.
- Fabra, Pompeu (1932), *Diccionari general de la llengua catalana*, Barcelona, Llibreria Catalònia.
- Fabra, Pompeu (1956), *Gramàtica catalana*, Barcelona, Teide.
- Fabra, Pompeu (2005–2013), *Obres completes*, edd. Jordi Mir/Joan Solà, 9 vol., Barcelona/València/Palma, Proa/3i4/Moll.
- Fabra, Pompeu (2009), *The Architect of Modern Catalan. Pompeu Fabra (1868–1948). Selected Writings*, ed. Joan Costa, Amsterdam/Philadelphia, Benjamins.
- Fayol, Michel/Jaffré, Jean-Pierre (2014), *L'orthographe*, Paris, PUF.
- Ginebra, Jordi/Solà, Joan (2007), *Pompeu Fabra: vida i obra*, Barcelona, Teide.
- Institut d'Estudis Catalans (1913), *Normes ortogràfiques*, Barcelona, Institut d'Estudis Catalans/L'Avenç.
- Institut d'Estudis Catalans (1997), *Documents normatius 1962–1996 (amb les novetats del diccionari)*, Barcelona, Institut d'Estudis Catalans.
- Institut d'Estudis Catalans (2007), *Diccionari de la llengua catalana*, Barcelona, Edicions 62/Enciclopèdia Catalana.
- Institut d'Estudis Catalans (2017), *Ortografia catalana*, Barcelona, Institut d'Estudis Catalans.
- Katz, Leonard/Frost, Ram (1992), *The Reading Process is Different for Different Orthographies: The Orthographic Depth Hypothesis*, Haskins Laboratories Status Report on Speech Research, SR 111–112, 147–160. Also as: *Reading in Different Orthographies: The Orthographic Depth Hypothesis*, in: Ram Frost/Leonard Katz (edd.) (1992), *Orthography, Phonology, Morphology and Meaning*, Amsterdam, Elsevier, 67–84.
- Kremnitz, Georg (2018), *Katalanische und okzitanische Renaissance. Ein Vergleich von 1800 bis heute*, Berlin/Boston, De Gruyter.
- Lafont, Robèrt (1971), *L'ortografia occitana. Sos principis*, Montpellier, Centre d'Estudis Occitans.
- Lamuela, Xavier (1991), *A Characterization of Alphabetical Writing Systems*, in: Utta von Gleich/Ekkehard Wolff (edd.), *Standardization of National Languages. Symposium on Language Standardization (Hamburg, 2–3 February 1991)*, Hamburg, Unesco-Institut für Pädagogik, 65–78.
- Lamuela, Xavier (1994), *Estandardització i establiment de les llengües*, Barcelona, Edicions 62.
- Lamuela, Xavier (1995), *Criteris de codificació i de compleció lingüístiques*, Els Marges 53, 15–30.
- Lamuela, Xavier (1996), *La codification du catalan au XXe siècle*, in: Henri Boyer (ed.), *Sociolinguistique: territoire et objets*, Lausanne, Delachaux et Niestlé, 159–177.
- Lamuela, Xavier (2017), *Une orthographe englobante pour le francoprovençal? Avantages et difficultés*, Nouvelles du Centre d'Études Francoprovençales René Willien 75, 68–98.
- Lamuela, Xavier/Murgades, Josep (1984), *Teoria de la llengua literària segons Fabra*, Barcelona, Quaderns Crema.
- Neustupný, Jiří V. (1970), *Basic Types of Treatment of Language Problems*, Linguistic Communications 1, 77–100.
- Recasens, Daniel (1991), *Fonètica descriptiva del català. Assaig de caracterització de la pronúncia del vocalisme i del consonantisme del català al segle XX*, Barcelona, Institut d'Estudis Catalans.
- Saussure, Ferdinand de (1972, 1916), *Cours de linguistique générale*, ed. Tullio De Mauro, Paris, Payot.
- Scharnhorst, Jürgen/Ising, Erika (edd.) (1976), *Grundlagen der Sprachkultur. Beiträge der Prager Linguistik zur Sprachtheorie und Sprachpflege*, 2 vol., Berlin, Akademie Verlag.

- Schmalz, Xenia, et al. (2015), *Getting to the Bottom of Orthographic Depth*, *Psychonomic Bulletin & Review* 22, 1614–1629. DOI: 10.3758/s13423-015-0835-2.
- Segarra, Mila (1985a), *Història de la normativa catalana*, Barcelona, Enciclopèdia Catalana.
- Segarra, Mila (1985b), *Història de l'ortografia catalana*, Barcelona, Empúries.
- Segarra, Mila (1991), *Pompeu Fabra*, Barcelona, Empúries.
- Vachek, Josef/Dušková, Libuše (edd.) (1983), *Praguiana: Some Basic and Less Known Aspects of the Prague Linguistic School*, Amsterdam/Philadelphia/Prague, Benjamins/Czechoslovak Academy of Sciences.
- Venezky, Richard L. (1977), *Principles for the Design of Practical Writing Systems*, in: Joshua A. Fishman (ed.), *Advances in the Creation and Revision of Writing Systems*, The Hague/Paris, Mouton, 37–54. Reproduced from: *Anthropological Linguistics* 12/7 (1970), 256–270.
- Venezky, Richard L. (1999), *The American Way of Spelling: The Structure and Origins of American English Orthography*, New York/London, The Guilford Press.
- Veny, Joan/Massanell, Mar (2015), *Dialectologia catalana. Aproximació pràctica als parlars catalans*, Barcelona/Alacant/València, Universitat de Barcelona/Universitat d'Alacant/Universitat de València.
- Vila, F. Xavier (2014), *Language Policy, Management and Planning*, in: Christiane Fäcke (ed.), *Manual of Language Acquisition*, Berlin/Boston, De Gruyter, 50–68.
- Wheeler, Max W./Yates, Alan/Dols, Nicolau (1999), *Catalan: A Comprehensive Grammar*, London/New York, Routledge.

Nicolau Dols

4 Phonology, Phonetics, Intonation

Abstract: This chapter offers a general overview of Catalan phonology, especially its segment inventory and the way segments combine into strings, and in so doing provides a description of syllabic patterns. Some phonetic detail is given in terms of the description of segments and also of basic intonation patterns. For descriptive purposes, the chapter also provides information on dialects, although this does not intend to be an exhaustive account of the various Catalan dialects, as the main focus of this chapter is phonological processes. These are classified as affecting prosody or the chronological tier of strings or as affecting melody or the feature content of segments. In either case, both vocalic and consonantal processes are analysed. Section 4 focuses on phonological processes triggered by morphological requirements. Specific attention is paid to phonology in conjugation. The final section presents the most typical suprasegmental patterns in stress assignment and intonation.

Keywords: vowels, consonants, syllables, morphophonology, intonation

1 Sound inventory

All sounds in Catalan are pulmonic, and on the whole follow predictable tendencies; there are no vowels marked for labiality (back vowels are rounded and non-back ones are unrounded), and there are no nasal vowels or trapezium-internal vowels other than /ə/. The places of articulation for consonants span from the lips back to the velum, i.e. there are no glottal consonants. Moreover, only obstruents can be unvoiced, and there is no opposition based on aspiration.

1.1 Vowels

Stressed vowels can be classified according to height and backness, assuming that roundness is an intrinsic feature of back vowels, as can be seen in Figure 1:

<https://doi.org/10.1515/9783110450408-005>

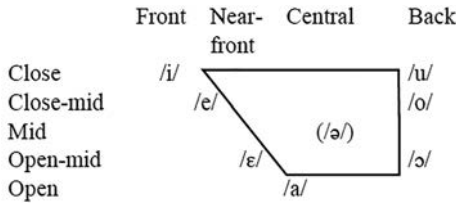


Figure 1: Vowel system.

The central mid vowel /ə/, deriving from the Latin *Ē/Ĭ*, as in *TRĒS* > [ˈtrəs], *VĪR(I)D(E)* > [ˈvərt] exists only in the Balearic dialects of Catalan when stressed, and, more generally, only in Eastern Catalan when unstressed. This presence as an unstressed vowel has become an axis of dialect classification.

According to Recasens i Vives (2014, 21), the referential values for the three first formants in Eastern Catalan are as follows:

	i	e	ɛ	a	ɔ	o	u
F1	334	450	581	730	608	489	394
F2	2078	1839	1700	1358	1125	1047	960
F3	2662	2571	2434	1125	2382	960	2431

The main differences affecting stressed vowels across dialects relate to (i) the presence of a stressed /ə/ in Balearic Catalan, as mentioned above, and (ii) a wider openness of open-mid vowels both in Balearic and Valencian Catalan. The formantic values that Recasens i Vives (2014) attributes to stressed /ə/ in Balearic are 563, 1393, and 2614 Hz for each of the first three formants. F1 for /ɛ/ is 659 for Balearic, and 601 for Valencian, and F1 for /ɔ/ is 708 for Balearic, and 621 for Valencian. Thus, both dialects exhibit wider mid-open vowels than Eastern (non-Balearic) dialects. The phonological opposition between close-mid and open-mid vowels has been elided in Northern Catalan, thus leading to a five-vowel system in which front mid vowels (/e/ and /ɛ/) and back mid vowels (/o/ and /ɔ/) merge into two vowels (near front /e/ and back /ɔ/).

As is to be expected, the unstressed vowel system is shorter than the stressed vowel system. Open and open-mid vowels (/ɛ, ɔ, a/) have no place in it, and the schwa (/ə/) appears in Eastern dialects to replace central and near-front vowels (/e, ɛ, a/). The exact substitutions involved in vowel reduction are analysed in 3.2.1.

1.2 Consonants

According to sonority, Catalan consonants span from unvoiced stops to glides. Only obstruents (stops, affricates and fricatives) produce voicedness oppositions. Sonorants can be grouped into nasals, laterals, rhotics and glides. Voiced approximants

also exist, though they do not belong to any inventory sound group, and only result from spirantisation, a process of weakening that affects voiced stops in the context of continuant sounds (to be seen in 3.2.2). As pointed out above, the velum is the furthest back point of articulation in Catalan.

Stops:

Bilabial		Dental		Velar	
Unvoiced	Voiced	Unvoiced	Voiced	Unvoiced	Voiced
p	b	t	d	k	g

A supplementary set of stops exists, palatal /c, ɟ/ alternating with velar /k, g/ before non-back vowels, and in word-final position in Majorcan Catalan. Dental consonants tend to be denti-alveolar pronounced with the apicolaminal region of tongue. However, pure dental stops are not uncommon, especially /d/. Velar stops range from postpalatal to velar proper depending on vowels and pauses surrounding the consonant. In the vicinity of pauses (previous or posterior) and front vowels, velars tend to become postpalatal (Recasens i Vives 2014, 57–87).

Voice Onset Timing measurements show a clear difference between voiced and unvoiced segments: negative values for voiced and low positive values for unvoiced discard aspiration (Julià 1981). Voice Offset Timing occurs before the beginning of a sentence-end stop, and thus shows no voicedness contrast in this position (see 3.2.2 on final devoicing).

Complex segments /k^w, g^w/ have been proposed (Lleó 1970; Wheeler 1979), to explain differences in syllabification, as in *quota* (bisyllabic) ['kwɔtə] 'fee' / *cuota* (trisyllabic) [ku'ɔtə] 'tail (augmentative)'.

Affricates:

Alveolar		Palatal	
Unvoiced	Voiced	Unvoiced	Voiced
ts̺	dz̺	tʃ̺	dʒ̺

Occurrences of /pf/ are only contextual and /bv/ appears only intramorphemically in *obvi* and derived words. These clusters are not considered to belong to the sound inventory as affricates. They do not alternate with other sounds and they do not appear in positions other than intervocalically. Voiced affricates do not appear at the beginning of words and unvoiced ones appear in that position only in loanwords (*tsunami*, *txadià*). Moreover, it is rare to see the unvoiced alveolar affricate (/ts̺/) in a morpheme-internal position, and, apart from loanwords (*jujitsu*, *matsutake*), it appears in only one case that is not attributable to compounding or loaning: *lletsó* (*pontsicà* deriving from *Ponts*, a town name derived from a plural; *quetsémper* [a kind

of fish] deriving from the Latin idiom *et nunc et semper*, and *sotsobrar* deriving from the prepositions *sots* and *sobre*) (see 3.1 on lengthening of affricates).

Fricatives:

Labiodental		Alveolar		Prepalatal	
Unvoiced	Voiced	Unvoiced	Voiced	Unvoiced	Voiced
f	v	s	z	ʃ	ʒ

The voiced labiodental fricative has been replaced by [b] ([β] when the spirantisation rule applies) in most dialects. It still remains in Balearic Catalan and in certain areas of Western Catalan, although it is at risk of being substituted in the speech of younger people. /s/ and /z/ are generally apicoalveolar, but they can be articulated by the tongue blade in the alveolar region, especially in Balearic and Algerese Catalan (Recasens i Vives 1991, 267). Voiced sibilants (/z/ and /ʒ/) can disappear in non-standard speech, namely in Central Valencian. /ʃ/ and /ʒ/ can be reinforced and appear as affricates in first position of a word and, especially, after a pause (see 3.2.2). Postvocalic prepalatal fricatives may appear accompanied by a preceding front glide (/j/) in Western dialects and in other scattered places (Pradilla Cardona 2002).

Nasals:

Bilabial	Alveolar	Palatal
m	n	ɲ

Other points of articulation can be found: labiodental [ɱ], dental [ɳ], prepalatal [ɲ] or velar [ŋ] with no contrasting function, as phonetic outcomes of /m, n/ followed by a consonant. /ɲ/ is pronounced with the tongue dorsum against the alveoli and the palate, and it is rare word-initially. Inside a word, it can only be followed by a final /s/.

Laterals:

Alveolar	Palatal
ɭ	ʎ

The alveolar lateral is normally pronounced with the tongue apex against the alveolar region or against the upper teeth (Bibiloni 2016), while the tongue dorsum is raised, without contact, towards the velum. This configuration is strongly adhered to in the Balearics (except among young speakers), and it is general in a word-final position or in the vicinity of back vowels. When two /ɭ/s appear in writing, a special symbol <·> is used to avoid confusion with the written form of /ʎ/ (<ll>), and, therefore, <1·1> is pronounced as a duplicated /ɭ/ (or as a single one if the written form is due only to

etymological reasons, as in *col·lisió*, *afil·le*), and <ll> is read as a palatal (*coll*, *callar*). /ʎ/ is pronounced with the tongue dorsum against alveoli and palate. Due to contact with Spanish and French, the palatal lateral can be heard (again in young speakers) as [j]. Nevertheless, some graphic <ll>s have been historically pronounced as [j] in Eastern dialects, a tendency maintained in the Balearics in the case of words deriving from Latin groups LJ (PALEA > *palla* ‘straw’), T’L (VETULUS > *vell* ‘old’), K’L (OCULUS > *ull* ‘eye’). In fact, this is just a case of a graphic symbol (<ll>) representing two different inventory sounds. Valencian and Balearic exhibit the graphic group <tl> (pronounced [ʎ]) in cases where other dialects have <tll> (pronounced [ʎʎ]) as in *espatla/espatlla*, *ametla/ametlla*.

Rhotics:

Flap	Trill
r	r

Both rhotics are apicoalveolar. They can hold phonological opposition only intervocally: *para* [ˈpaɾə] ‘(s/he) sets (as in *s/he sets the table*)’/ *parra* [ˈpaɾə] ‘vine’. The only difference between a flap and a trill is in the number of contacts between the tongue tip and the alveoli, only one for a flap and between two and four for a trill. In word-final position, if the rhotic does not simply disappear (see 4), it can be produced as a flap or a trill, depending on the dialect, with no phonological significance, except for verbs in 1st pers. sg. in Balearic, where no morpheme is added to the stem: *amar* [əˈmar] ‘I tie’/ *amar* [əˈmar] ‘I soak’ (Dols/Wheeler 1996). In word- or stem-initial position, only [r] appears (*romà* [roˈma] ‘Roman’, *preromà* [preroˈma] ‘pre-Roman’), and it is also the only possible rhotic in the syllable onset following a syllable-final consonant (*folre* [ˈfoɫrə] ‘lining’), whereas [r] is the only rhotic to follow a homosyllabic consonant: *tros* [ˈtrɔs] ‘piece’, *frare* [ˈfrarə] ‘friar’. In preconsonantal position, rhotics can be pronounced as flaps or lighter trills (as a double flap, Wheeler 2005), also depending on specific dialects, but again, as in word-final position, no phonological opposition remains. Depending on the tempo, the rhotic flap can be replaced by an alveolar approximant ([ɹ]) intervocally. In Northern Catalan, due to the influence of French, the dorsouvular rhotic ([ʀ]) and dorsouvular fricative ([ʁ]) can also be heard.

Glides:

Palatal	Velar
j	w

Graphically rendered as <i, u>, glides are rare in word-initial position, in most cases due to loanwords (*iot* [ˈjɔt] ‘yacht’, *uacari* [ˈwakari] ‘uakari’). They can be in syllable-

initial (*cauen* ['kawən] 'they fall', *jaia* ['zajə] 'old woman') and syllable (and word) final position (*taula* ['tawlə] 'table', *mai* ['maj] 'never'), but they are often blocked when following a homosyllabic consonant due to a tendency to avoid raising diphthongs after a consonant. However, following a sibilant after the word stress or in stress position in certain suffixes, this blocking effect seems to disappear, as in *gràcia* ['grasjə] or *oració* [orə'sjo] 'sentence'. In other cases, the corresponding close vowels appear: *viatge* [vi'adʒ:ə] 'journey', *cruel* [kru'ɛɫ] 'cruel'. [j] and [w] can be the result of sandhi affecting an initial or final close vowel in contact with a preceding or following vowel: *fa impressió* [fajmprə'sjo] 'it impresses', *té una casa* [tewnə'kazə] 's/he has a house'.

2 Syllabic structures

Syllables in Catalan can contain a coda and lack an onset. Therefore, the basic patterns are those resulting from the schema (Onset) – Nucleus – (Coda). Dols/Wheeler (1996) point also to syllables with no nucleus to explain counter-sonicity finals in 1st pers. sg. present indicative in Balearic Catalan (see 4.3 below on verb phonology). Permitted contents in syllable constituents are explained in the following two sections.

2.1 Nuclei

The syllable nucleus always contains a vowel, and no other segment type is allowed in this position. All vowels in the inventory can occupy the nucleus of a stressed syllable, except for /ə/ in dialects other than Balearic. For this reason, its status as an inventory sound in the other dialects is not clear, and this is a topic of discussion in taxonomic phonology (reviewed in Mascaró 1991).

Much speculation remains surrounding the feasibility of diphthongs in Catalan. If a diphthong is “a complex vowel of non-steady quality, made up of two phases” (Roca/Johnson 1999, 688), or a “a phonetic sequence, consisting of a vowel and a glide, that is interpreted as a single vowel” (SIL “diphthong”), and, based on these assumptions, both parts of a diphthong should be encompassed in the syllable nucleus, then it should be concluded that diphthongs are not possible in Catalan, at least according to the classic descriptions of the language. If a diphthong is no more than a shorthand notation for the sequence of a glide and a vowel or a vowel and a glide regardless of whether they are assigned to the syllable nucleus or to the margins, then the existence of diphthongs, with a heavy preference for falling diphthongs (vowel-glide) over raising (glide-vowel), is plausible. The main argument against the existence of nuclear diphthongs in Catalan is the presence of certain length restrictions on margins (a test already put forward by Selkirk 1982 for English diphthongs); a

complex syllable onset is not possible if followed by a raising diphthong, as in the following examples:

tri.bu ['tri.βu] 'tribe', *tri.ar* [tri.'a] 'to choose' **triar* ['trja]

A similar test shows how restrictions on coda length precludes the consideration of vowel-glide sequences as wholly dependent on the nucleus node; there are no words with a longer final syllable due to the inclusion of /j/ or /w/ hypothetically assigned to the nucleus to overcome restrictions on codas. A word such as *aire* ['ajrə] has a final epenthetic vowel triggered by the ill-formed /ajr/. Only the words *rail*, *saur*, *vair* and *cuir* exhibit a glide followed by a liquid at the end of the word (Wheeler 1987), and they all have alternative forms in Catalan, namely *rail* (the only normative form after the 2016 reform of orthography), *saure*, *vaire* and *cuiro*, all of them avoiding a final sequence which would be fully acceptable should the glide belong to the syllable nucleus. Apart from this, in dialects with stronger limits on consonantal strings (Majorcan and Minorcan) it is clear that postvocalic glides count as coda consonants for string simplification as in *be.neit* [bə'nəjt] 'silly (masc. sg.)' / *be.neits* [bə'nəjs] (*[bə'nəjts]) 'silly (masc. pl.)'.

Regardless of their classification as diphthongs or as consonantal clusters, sequences formed by a glide and a vowel are fully accepted when in word-initial position (although this is rare) or after a preceding vowel (as already stated in "Glides", 1.2 above): *iot*, *uacari*, *jaia*, *cauen*. In postconsonantal position, the sequence has been traditionally discarded as a product of Spanish influence (Badia i Margarit 1962), and the modern standardisation of Catalan spelling systematically splits such sequences into two different syllables when classifying them for graphic accent purposes: *gràcia* is analysed as *grà.ci.a* and classified as a paroxytone, and, likewise, *tènue* 'light', *àrdua* 'hard', *història* 'history', etc. The issue has become a common point of interest in discussions on scansion in Catalan (Alcover 1908; Coromines 1971; Cardona 1977; Oliva 1980; 1988; 1992; Recasens i Vives 1990; Bargalló Valls 2007). An argument that appeared as early as in Fabra (1891) and runs right up to the recent official grammar of Catalan (GIEC 2016) contends that not all positions in the word should be analysed in the same terms; in particular, post-tonic sequences following a sibilant consonant tend more naturally towards agglutination than towards diaeresis. According to Vallverdú Albornà (2002), this tendency applies to any word-final group and to certain tonic groups beginning with /i/, from which those following /r/ are categorically excluded, as in *Sarrià* '(place name)' [sə'ri'a]/*[sə'rja]. Vowel contacts between two words may depend on stress issues. Clitics like the masculine definite article (*el*), the preposition *de* or pronominal clitics (*em*, *et*, *el*, *ens*, *els*, *en*) contain vowels not present in underlying forms and, thus, these only emerge due to syllabification needs as in *l'assumpte del teu company m'amoïna/el problema del teu company em preocupa* 'your colleague's problem worries me', where no contact between vowels belonging to different words occurs because, in the first sentence, there is no need for epenthesis. In other cases, where vowels actually arise in speech, resyllabification can

unite two vowels originally belonging to two different words: *ara i sempre* [arəj'semprə] 'now and ever', *lliçó unificada* [li,sownifi'kaðə] 'unified version'. There seems to be a restriction on this sort of agglutination: it is normally avoided when it would produce a clash between two stressed syllables; *cançó inèdita* [kən,soi'nɛðitə] is more common than [kən,soj,nɛðitə], where a clash would result (Oliva 1992, 160–161; Wheeler 2005, 127). A reaction in the opposite direction, i.e. diphthongisation between word boundaries intended to avoid a metrical lapse, is not as strongly adhered to in order to avoid clashes: *raó increïble* [rə,ojŋkrə'ibblə] is not imposed onto [rə,ojŋkrə'ibblə] for that reason, and both solutions are possible depending on speech tempo, which in fact may also affect clash avoidance since this is not an indispensable requirement in Catalan. At the end of the scale of acceptability, strongly agglutinated pronunciations such as [aɛ] for the contact between *portar* and *herbes* in *va portar herbes amargues* (Vallverdú Albornà 2002, 143) can occur provided that speech tempo is high and the sequence does not hold the main phrase stress. Observe that in this latter case it is not a close vowel which loses syllabicity, but a mid-close one. Similar examples are given with an [ɔ], as for *va contemplar ones a la platja* (Vallverdú Albornà 2002).

2.2 Margins: onset and coda clusters

All consonants can be associated to both syllable margins, although some restrictions exist on word-boundary and word-internal syllables. The rhotic flap (/ɾ/) cannot appear in word-initial syllable onsets, and the voiced alveolar sibilant (/z/) is rare, only appearing in learned words, as in *zona*, *zebra*, *zoo*. Palatal full consonants (/ʎ, ɲ, ʝ, ʒ, tʃ, dʒ/) do not appear in word-internal codas, which is to say that they cannot appear in a preconsonantal position. This also holds true for sibilants in pre-word-final position, but not for /ʎ, ɲ/, which can be followed by a word-final /s/ as in plurals or in 2nd pers. sg. verb forms: *balls* ['baʎs] 'dances', *empenys* [əm'pɛɲs] 'you push'. The restriction against palatal full consonants in word-internal codas is not applied to the palatal glide (/j/): *gaire* ['gajrə] 'much', *afaitar* [əfəj'ta] 'to shave'.

Syllable margins can be complex, i.e. containing more than one consonant. Complex margins follow the sonority scale (Kiparsky 1979; Selkirk 1982) as in the sequences listed below.

Complex onsets: any plosive (except for complex /kʷ/ or /gʷ/ mentioned above in 1.2) or /f/ can be followed by a flap (/ɾ/) or an alveolar lateral (/ɭ/) (except for the groups */tɭ/, */dɭ/)

	ɾ	ɭ
p	prou 'enough', comprar 'to buy'	plaça 'square', omplir 'to buy'
b	brusc 'rough', obrir 'to open'	blat 'wheat', cable 'wire'

t	tros ‘bit’, quatre ‘four’	
d	drac ‘dragon’, adreça ‘address’	
k	creu ‘cross’, escriure ‘to write’	clau ‘key’, oncle ‘uncle’
g	gra ‘grain’, sogre ‘father-in-law’	glaçó ‘ice cube’, regla ‘rule’
f	frare ‘friar’, cofre ‘chest’	flor ‘flower’, inflar ‘to swell’

Longer onsets are not permitted phonetically, although initial /s/ followed by a consonant or a consonantal cluster can be on occasions easily understood as belonging to the underlying representation due to alternations such as the following:

escriure [əs'kriwrə] ‘to write’ – *inscriure* (with prefix *in-*) [ins'kriwrə] ‘to register’ – *proscriure* (with prefix *pro-*) [pros'kriwrə] ‘to outlaw’
estrènyer [əs'trɛɲə] ‘to narrow’ – *restrènyer* [rəs'trɛɲə] (with prefix *re-*) ‘to constipate’ – *constrènyer* [kons'trɛɲə] (with prefix *con-*) ‘to constrain’
estar [əs'ta] ‘to be’ – *constar* [kons'ta] (with prefix *con-*) ‘to consist of’ – *instar* [ins'ta] (with prefix *in-*) ‘to urge’

In cases where the stem aligns with the beginning of the word, with no affix preceding it, epenthesis provides the vowel [ə] (or [e] in Western dialects, often pronounced [a] in initial closed syllables), as in the three first examples above (*escriure*, *estrènyer*, *estar*).

Complex codas are not as straightforward. Final clusters regulated by the sonority scale – now implying a decrease (a mirror image of sonority in the onset) – are more numerous than onset clusters. It should be noted that (a) due to the devoicing rule, in actual pronunciation plosives and fricatives do not maintain voicedness oppositions in word-final position (see 3.2.2) and, therefore, many of the cases below only bear witness to the lexical origin of clusters, but not actual pronunciation: *calb* is pronounced [ˈkaɫp], and (b) some of the examples below are learned or loan words that occur fairly infrequently such as, for instance, all those with a penultimate glide.

		Penultimate consonant			
		Fricative	Nasal	Liquid	Glide
Last consonant	Plosive	sp, st, sk, sg (1)	mp, nt, nk, mb, nd, ng (2)	lp, lt, lk, lb, ld, rp, rt, rk, rb, rd, rg (3)	jp, jt, jk, jd, wt, wk (4)
	Fricative		mf, ns, nj (5)	ls, lf, rs, rf, rf (6)	js, ws (7)
	Nasal			lm, rm, rn (8)	jn, wm, wn (9)
	Liquid				jr, jl, wr (10)

Examples: (1) fricative-plosive: *gesp* ‘(a species of grass)’, *capitost* ‘leader’, *bosc* ‘wood’, *pelasg* ‘Pelasgian’; (2) nasal-plosive: *camp* ‘field’, *cant* ‘chant’, *banc* ‘bank’, *tomb* ‘turn’, *fecund* ‘fertile’, *sang* ‘blood’; (3) liquid-plosive: *talp* ‘mole’, *alt* ‘tall’, *solc*

‘furrow’, *calb* ‘bald’, *herald* ‘herald’, *esquerp* ‘elusive’, *curt* ‘short’, *corc* ‘woodworm’, *corb* ‘raven’, *tard* ‘late’, *amarg* ‘bitter’; (4) glide-plosive: *naip* ‘playing card’, *beneit* ‘silly’, *laic* ‘layman’, *alcaid* ‘warden’, *caut* ‘cautious’, *rauc* ‘harsh’ (voice); (5) nasal-fricative: *triomf* ‘triumph’, *ens* ‘being’, *ponx* ‘punch’ (drink); (6) liquid-fricative: *pols* ‘dust’, *escalf* ‘warmth’, *cors* ‘Corsican’, *serf* ‘serf’, *guerx* ‘warped’; (7) glide-fricative: *rais* ‘rais’ (Arabian title), *bordeus* ‘Bordeaux’ (wine); (8) liquid-nasal: *calm* ‘calm’, *erm* ‘barren’, *carn* ‘meat’; (9) glide-nasal: *ain* ‘ain’ (Arabic, Hebrew and Aramaic letter), *linòleum* ‘linoleum’, *clown* ‘clown’; (10) glide-liquid: *cuir* ‘leather’, *gasoil* ‘diesel oil’, *saur* ‘dark yellow’.

All these clusters above can even be enlarged through adjunction of a final morphemic /s/, except if the original cluster already ends with a sibilant, in which case epenthesis or invariability applies. However, the original cluster or the enlarged one can be simplified by elision in different extents according to dialects (*tomb* is [tom] in Central, Northwestern, Northern, and Eivissenc Catalan, but [tomp] in Majorcan, Minorcan, Valencian and Algherese Catalan; see 3.1).

Word-internally, complex codas are limited to the clusters formed by consonant-/s/, or /mp/, /ng/, often built through derivation or compounding: *trans.criu.re* ‘to transcribe’, *subs.ti.tut* ‘substitute’, *ex.pli.car* [əks.pli.ka], *comp.te* ‘bill’, *sang.tra.ït* ‘bruise’. Cluster simplification is common in careless speech (see 3.1).

3 Phonological processes

3.1 On the prosodic tier: elision, epenthesis, lengthening, split

Strategies for eliminating sequences of syllable nuclei are not unusual across languages (Jakobson/Halle 1956; Vennemann 1988; Prince/Smolensky 1993). As seen above, this can trigger resyllabification of close vowels as glides (see 2.1). Another way to conflate syllables whose vowels are in contact is by deleting one of them. The general rule for vowel deletion relates to phonological prominence: unstressed vowels tend to disappear if (1) they are identical to a neighbouring vowel, or (2) one of the vowels is [ə], the most unmarked unstressed vowel and, therefore, often the result of epenthesis. The study of vowel elision in Catalan stands at the crossroads between phonological legitimacy relating to the axis underlying/epenthetic, and the ruling role of prosody in the organisation of the phonetic representation. If we start with the examples in (1) and (2), the importance of factors such as distance from the phrase stress or a close syllable in second position easily stand out:

- (1) *porti informes* [pɔr.tim̩.foɾ.məs] ‘bring (imperative, 2nd pers. sg., polite) reports’, *deixo obertures* [de.fu.βɛr.tu.rəs] ‘I leave gaps’, but *doni idees* [do.ni.i.ðe.əs] ‘give (imperative, 2nd pers. sg., polite) ideas’, *antiinflamatori* [aɲ.tim̩.flə.mə.ˈto.ri] ‘anti-inflammatory’, *portaavions* [pɔr.tə.vi.ˈons] ‘aircraft carrier’, but *microones* [mi.kro.ˈo.nəs] (in non vowel-reduction dialects) ‘microwaves’, *xiisme* [ʃi.iz.mə] ‘Shiism’.

- (2) *entendre això* [əŋ.ten.drə.'ʃə] 'to understand this', *aquesta illa* [ə.kes.'ti.ʎə] 'this island', *última hora* [uʎ.ti.'mɔ.rə] 'last hour', *somni estrany* [som.nis.'trən], but *hindú asmàtic* [iŋ.du.əz.'ma.tik] 'asmatic Hindu', *somni eròtic* [sɔm.ni.ə'.rɔ.tik] 'erotic dream'.

The last two examples in (1) show identical word-internal vowels, except for the stress on the second one. The last two examples in (2) have been proposed by Palmada (1994, 123) to show how an initial /ə/ never undergoes elision when in contact with a preceding vowel, as it can be maintained that [ə] in *estrany* and other words not alternating with words containing a stressed vowel in the same place is epenthetic and thus does not emerge in continuous syllabification of a sound string where an initial consonant can be abutted by a preceding word-final vowel. On the other hand, the [ə] in *eròtic* cannot be epenthetic, for it is the only sound in its syllable.

The relation between elision and stress has been thoroughly analysed in Wheeler (2005). Both clashes and lapses arising due to vowel elision would prevent the process from having effect: *destí idoni* 'suitable fate' is better with no elision ([dəs.ti.i.'ðɔ.ni]) than with an elided vowel (*[dəs.ti.'ðɔ.ni]), a pronunciation that would include a stress clash. Conversely, *hindú honradíssim* 'most honest Hindu' is preferable with vowel elision than without in order to avoid too long a lapse between stresses: [iŋ.dun.rə.'ðɪ.sim] is better than *[iŋ.du.un.rə.'ðɪ.sim].

Beyond factors related to phonological legitimacy and prosodic structure, other examples point toward the importance of morphological domains: in *tu obriràs* 'you will open' or *pi inclinat* 'bent pine' elision would affect not only a vowel cluster between a final and initial syllable, but also between two initial syllables, and word recognition could suffer from this. The preferred pronunciations are therefore [tu.u.βri.'ras], *pi inclinat* [pi.iŋ.khi.'nat], although avoidance of excessive lapses can force elision as in *pi inclinadíssim* 'most bent pine' [piŋ.khi.na.'ðɪsim]. However, it is not easy to provide clear-cut rules in this area of phonological competence, and much should be attributed to speech tempo and care taken when speaking, with diverging results being easy to find in real life contexts.

In colloquial Catalan the elision of /ə/ between a stop (especially labial) or an /f/ and rhotic flap is common, especially in the Central dialect, and it creates a complex syllable onset as those seen above in 2.2. Examples of syncope are *barana* ['branə] (standard [bə'ranə]) 'railing', *taronja* ['trɔŋzə] (standard ['təɾɔŋzə]) 'orange', *safareig* [sə'fɾeɪʃ] (standard [səfə'reɪʃ]) 'pond'.

Two instances of consonantal elision deserve mention: those in word-final and word-internal position. Two instances of single consonants elided in word-final position are well known in the phonology of Catalan: /-r/ and /-n/. Since they have a morphological trigger, these two will be addressed to in 4.1.

Word-final consonant clusters ending in a stop tend to disappear in Central, Northwestern, Northern and Eivissenc, and tend to persist in Valencian, Majorcan, Minorcan and Alguerese. Dialects simplifying such clusters elide the final stop when

its point of articulation resembles that of a preceding consonant as in (3) but tend to maintain the cluster when the two consonants disagree in the point of articulation, as in (4). When coincidence between consonants is partial, elision appears as an option (5):

- (3) *rum* ['rum] 'course', *tant* ['tan] 'so much', *alt* ['aɫ] 'tall', *sang* ['san] 'blood': the stop can reappear optionally in *-ng* clusters when followed by an initial vowel as in *sang i fetge* – literally 'blood and liver', figurative for 'bloodshed' – (Bonet/Lloret 1998, 114).
- (4) *calb* ['kaɫp] 'bald', *solc* ['soɫk] 'furrow', *cresp* ['kɾɛsp] 'rippled', *tosc* ['tosk] 'rough', *corb* ['kɔɾp] 'raven', *corc* ['kɔɾk] 'woodworm'.
- (5) *post* ['pɔst]/['pɔs] 'board', *tard* ['taɾt]/['taɾ] 'late'.

The clusters in (4) and (5) exhibit a tendency to maintain both consonants before an additional [s], except for *-sts*, where /t/ is elided. *-st* and *-sp* can be abutted by an epenthetic vowel in masculine plurals (see 4.2 below).

A different sort of word-final elision affects stops when a following word commences also with a stop. This occurs in Southern Valencian, and colloquially in all of the other dialects, and it is more frequent with a word-final dental than with other stops, which can undergo elision if the following stop shares its point of articulation: *tot bo* [to'βɔ] 'all good', *ho sap bé* [usa'βe] 's/he knows it well', *sac gros* [sa'ɣɾɔs] 'big sack'.

Majorcan and Minorcan Catalan show a strong reluctance against word-internal consonant clusters. The basic rule is: "Allow three-consonant clusters only when a syllable border lays between the first and the second. Simplify all other clusters until they become acceptable" (Dols 2000, 328–329). Examples of acceptable consonantal clusters appear in (6), while (7) shows cases of repaired excessive strings:

- (6) *entrada* [əɲ.ˈtra.ðə] 'entrance', *altre* ['aɫ.trə] 'other', *ungla* ['uɲ.glə] 'nail', *inflació* [im.fə.si.ˈo] 'inflation'.
- (7) *instruir* [is.tru.ˈi] 'to instruct', *obstar* [os.ˈta]/[us.ˈta] 'to preclude', *superstició* [su.pər.ti.si.ˈo] 'superstition'.

The same can be said for these dialects when the cluster is created by word contact, as can be seen in *camp sembrat* [ˌkan.səm.ˈbrat] 'sown field' or *porc magre* [ˌpɔɾ.ˈma.ɣɾə] 'thin pig'. Only Minorcan maintains three-consonantal clusters where the first consonant is a liquid or a glide, as in *vols caure* 'you want to fall', *fers molt* 'you hurt a lot', *caus molt* 'you often fall' (Pons Moll 2007, 319), and the exact resolution can differ between both dialects, as in (8):

- (8) *tens por* [tem.ˈpɔ] (Maj.)/[tes.ˈpɔ] (Min.) 'you are afraid'; *pocs dies* [pɔd.ˈdiəs] (Maj.) / [pɔz.ˈðiəs] (Min.) 'few days'.

Some similar elisions are reported in other dialects, such as of the final /s/ of clitics in Gironese Catalan when followed by an initial consonant (*els dos* [əɫˈdos] 'the two (of

them)', *uns casos* [uŋ'kazus] 'some cases'), or more generally in careless speech as in labial sequences, as in *triomf bèl·lic*, or when /s/ stands between two consonants as in *exclamar* [əsklə'ma], *arcs gòtics* [arʒ'ɣɔtiks] (Recasens i Vives 1993, 197–198).

Vowel epenthesis has already been alluded to as a solution for ill-formed syllable margins and also as an explanation for alternating forms with and without the vowel [ə] in the discussion on vowel elision. A morphological case of epenthesis will be presented in 4.2, and in 4.3 the blocking of epenthesis due to morphological reasons will be examined in detail. Traces of schwa or [e] can be heard in Valencian and Alguerese after a phrase-final stop due to the strength of the plosive's release.

Consonantal epenthesis is not frequent in Catalan. However, a special instance of it will be analysed in 4.3 below. Some irregular or dialect-particular cases of such a process include the insertion of final [t], and the insertion of intervocalic [j] and [v].

The insertion of final [t] can occur after oxytone words ending in a rhotic in Northern and Central Catalan, as in *cor* ['kɔrt]/['kɔrt] 'heart', *amor* [ə'mɔrt]/[ə'mɔrt] 'love'. [t] may also be inserted after a final /i/ in paroxytone words with an open or a mid-open stressed vowel /a, ɛ, ɔ/: *api* ['apit] 'celery', *geni* ['ʒɛnit] 'genius', *premi* ['premit] 'prize', *somni* ['sɔmit] 'dream' (Bibiloni 2002, 284). This phenomenon is blocked in conjugation, where the 1st and 3rd persons of present subjunctive (ending in /i/) never experience this type of epenthesis. Oliva/Serra (2002) relate this epenthesis to the effects of preaccented suffixes (see 5.1 below).

Sequences of two non-close vowels can be split by the insertion of a [j]: *idea* [i'ðejə], *teatre* [tə'jatɾə], *paella* [pə'jeʎə]. This process appears in colloquial Central and Northwestern Catalan and in Eivissenc. In Northern Catalan it can affect sequences with an /i/: *ahir* [ə'jirə] 'yesterday', *aïnes* [ə'jinəs] 'tools'.

Between two vowels, of which one is rounded, a [v] can appear in Balearic, more often in Majorcan Catalan: *coa* ['kovə] 'tail', *lleó* [ʎə'vo] 'lion', *coent* [ko'vent] 'spicy hot'.

Strengthening processes include affrication and lengthening. The affrication of the voiced prepalatal fricative (/ʒ/) in word-final position is general, and its voicedness depends on general rules of devoicing or voicedness assimilation: *rajar* [rə'ʒa] 'to stream' / *raig* | ['ratʃ] 'stream' – *raig intens* [ˌradʒːiŋ'tɛns], *vegi* ['bɛʒi] 'I see (subjunctive)' / *veig* || ['bɛtʃ] 'I see (indicative)' / *veig això* [ˌbɛdʒːə'fɔ] 'I see this'. An overuse of this affrication can be found in several dialects, where it also appears intervocalically. It can be found in varying degrees in Valencian (in dialects where no j-split or devoicing occurs), in Northwestern Catalan and in Balearic: *roja* [ˌrɔdʒːa] 'red-brown (fem. sg.)', *llegir* [ʎəɖːdʒi] 'to read'. In Central Valencian, in the town of Gandia, and in the north of the Western Strip (known in Catalan as "la Franja") affrication and devoicing of /ʒ/ is constant in all positions and occurs together with devoicing of the alveolar fricative (/z/ → [s]) (Pradilla Cardona 2002, 310–312).

Affricates can be lengthened (i.e. pronounced with a longer initial occlusion) in some circumstances. In Central Catalan this process occurs intervocalically and preferably in post-tonic position. Lengthening seems more constant in voiced than

unvoiced affricates. According to Bonet/Lloret (1998, 177–178), underlying voiced affricates are always long intervocalically, whereas unvoiced ones are lengthened only when intervocalic and posttonic: *jutge* [ˈʒud͡ʒːə] ‘judge’/ *jutgessa* [ʒud͡ʒɛsə] ‘female judge’, but *despatxar* [dəspət͡ʃa] ‘to dispatch’/ *despatxa* [dəsˈpat͡ʃa]. In Majorcan and Minorcan Catalan there is a general tendency towards affricate lengthening in all cases where the affricate is in the intervocalic position. Affricates are shorter in Valencian. In other dialects, length can depend on the underlying form of the segments (Recasens i Vives 2014, 273).

Voiced stops /b/ and /g/ can be lengthened when followed by /l/ intervocalically: *doble* [ˈdobːlə] ‘double’, *segle* [ˈsegːlə] ‘century’. Bonet/Lloret (1998, 93–96) point at the fact that words such as *aglà* ‘acorn’ or *ègloga* ‘eclogue’ do not exhibit the same behaviour and conclude that those stops are lengthened before a morpheme-final /l/, an explanation consistent with an observation by Mascaró (1987) on coda consonants and the assumption that before final-vowel epenthesis the stop might have been in that syllable position. This process is strongly adhered to in Balearic and Northern Catalan and is also present in Central and Northwestern Catalan.

Palatal consonants show a tendency to split. This occurs with sibilants /ʒ/ and /ʃ/, and also with nasal /ɲ/. In all cases the split consists of a palatal glide [j] detaching from the original sound, which can remain palatal or lose its point of articulation. In Northwestern Catalan (especially in the south of the region, Pradilla Cardona 2002, 306–309), intervocalic /ʒ/ becomes [jʒ]: *roja* [ˈrɔjʒa] ‘red-brown (fem. sg.)’, *vagi* [ˈvajʒi] ‘I go (subjunctive)’. In some places this phenomenon can also affect affricates, as in *formatge* [forˈmajʒe] ‘cheese’. In Western dialects and in areas of Northern and Central Catalan, although not necessarily linked to palatal split in voiced sibilants, detachment can affect unvoiced ones both intervocalically and in word-final position, in a similar way to spelling, as in *caixa*, *això*, etc. In Valencian this detachment may imply depalatalisation of the sibilant, as in [ˈkajʃa], [ajˈʃɔ].

In Majorcan and Minorcan Catalan, any /ɲ/ followed by a consonant is divided between its palatal and nasal features, the first as a glide ([j]) and the second as a nasal taking the point of articulation of the following consonant: *l’any passat* [ləjnpəˈsat] ‘last year’, *codonys* [koˈðojns] ‘quinces’. In Majorcan Catalan this process can also be triggered by the adjunction of a consonant to a cluster [ɲc] / [ɲj] (see 1.2 above on these consonants) (Mascaró 1986): *troncs* [ˈtrojns].

3.2 On the melodic tier

3.2.1 Vowels: reduction and harmony

Vowel reduction relates to the difference between the stressed system and the unstressed set of vowels. Owing to the smaller size of the latter, whenever a syllable loses stress in favour of another on its right side due to a morphological process, its

vowel can change quality if it is not included in the unstressed vowel inventory. The unstressed vowel inventory can differ according to dialects. These are the most usual:

- (9) 1. Western: /i, e, a, o, u/.
2. Eastern (except Majorcan): /i, (e), ə, (o), u/.
3. Majorcan: /i, (e), ə, o, u/.

The parentheses in (9) indicate segments considered to be exceptions to the main rule. In these inventories open-mid vowels /ɛ/ and /ɔ/ are not found in the unstressed vowel system. Apart from this, unstressed /a/ is replaced by [ə] in all Eastern dialects, where /o/ is generally replaced by [u] apart from in the case of Majorcan and lexical exceptions in the other dialects. Relation rules are easy to understand bearing in mind that [i] reflects only underlying /i/, i.e. it does not reflect any other originally stressed vowel: *pi* [ˈpi] ‘pine’ → *pinassa* [piˈnasə] ‘pine needles’. In (9.2) [u] reflects all original round vowels: *bo* [ˈbɔ] ‘good’ → *bondat* [buɲˈdat] ‘goodness’, *tot* [ˈtot] ‘everything’ → *total* [tuˈtaɫ] ‘total’, while in (9.1) and (9.3) unstressed [o] reflects /o/ and /ɔ/, and [u] stands only for /u/: *bo* [ˈbɔ] → *bondat* [boɲˈdat], *tot* [ˈtot] ‘everything’ → *total* [toˈtaɫ] ‘total’, *lluny* [ˈluɲ] ‘far’ → *llunyà* [luˈɲa] ‘distant’. /ɛ/ is replaced by [e] in (9.1), while [ə] stands for /e/, /ɛ/, and /a/ in (9.2) and (9.3): *crèdit* [ˈkreðit] ‘credit’ → *acreditat* [akreðiˈtat] (9.1) / [əkɾəðiˈtat] (9.2), (9.3) ‘accredited’. In alguerese, and partially in Barcelona Catalan, /e/ and /ɛ/ can be replaced by an unstressed [a] (Mascaró 2002).

Exceptions to vowel reduction can depend on phonetic, morphological or lexical reasons. In the Eastern dialects, contact of a vowel tending to [ə] with a previous or a following vowel can block reduction, as in *teatre* [teˈatrə] ‘theatre’ or *aeroport* [aeruˈport] ‘airport’ (Recasens i Vives 1993; Mascaró 2002). Reduction avoidance seems to be stronger when the vowel in question is preceded by another one. In prevocalic position the exception rule is less than clear. In some Western dialects, [ɔ] and [ɛ] can appear in unstressed syllables when spread by vowel harmony (see below). Compounds and derived words bear a secondary stress that can protect the vowel’s integrity. This explains why manner adverbs ending in *-ment* do not show reduction in the stem (see 5.1 below): *bonament* [ˌbɔnəˈmen] ‘in good manners’, *fàcilment* [ˌfasitˈmen] ‘easily’, *francocatalà* [ˌfran̩kukətəˈla] ‘French-Catalan’. A high degree of compound lexicalisation can be observed in cases of reduction affecting an originally stressed element, as in *potser* [puˈt̪sɛ] in Central Catalan. Derived words in Majorcan Catalan, especially augmentatives, diminutives and verbs, show a tendency to avoid complete reduction to [ə] if the stem’s vowel is /ɛ/ or /e/: *llet* [ˈɫet] ‘milk’ → *lleteta* [ɫeˈtətə] ‘milk’ (diminutive), but *lletera* [ɫəˈterə] ‘milk pot’, *esper* [əsˈper] (‘I wait’) → *esperar* [əspeˈra] ‘to wait’. In verbs this avoidance affects mainly vowels in a palatal context (*aixecar*, *engegar*), although it can also occur in other cases, as in the example above. The rule is elusive and the blocking appears idiosyncratic. In Central Catalan round vowels followed by [w] show dissimilation when they lose their stress in order to avoid a poorly perceivable [uw]. In such cases, [əw] appears: *plou* [ˈpləw] ‘it rains’

→ *plourà* [pləw'ra] 'it will rain', *roure* ['rowrə] 'oak' → *roureda* [rəw'reðə] 'oak wood'. Other cases of reduction that are avoided are lexical: acronyms, loans or learned words, as in *UNESCO* [u'nesko], *salve* ['saɫβe], *Boston* ['boston].

Instances of vowel harmony exist in Catalan, both progressive and regressive. In progressive (i.e. rightwards) harmony, an unstressed /a/ can copy a preceding [ɛ], or [ɔ] in Valencian towns South of river Xúquer: *terra* ['tɛrɛ], *roca* ['rɔkɔ]. Harmony is not uniformly distributed in the area (see Colomina i Castanver 1985). According to Montoya (1981) total harmony (i.e. rightwards and leftwards simultaneously) is observed in cases such as *efecte* [ɛ'fɛkte] 'effect', *tovallola* [tɔvɔ'ɔɫɔ] 'towel'. In Jiménez (1998) instances of unstressed final [ɔ] in cases where no harmony can be supposed, such as in *mira* ['mirɔ] 'look' (observed in Ontinyent), are explained as analogies to words undergoing harmony.

In parts of Majorcan and Northwestern Catalan, progressive harmony can trigger closure of an unstressed /o/ to [u] by action of a following stressed /i/ or /u/, as in *molí* [mu'hi] 'windmill', *comú* [ku'mu] 'common'. Dols (2011) shows how also a pretonic /i/ can trigger this process, as in Majorcan *cossiòl* [kusi'ɔɫ] 'flower pot'.

3.2.2 Consonants: feature-changing phenomena

Consonants can only maintain their voicedness quality in onsets; when in coda, they take the same voicedness as a following consonant (resyllabification places any prevocalic consonant in an onset). This rule also applies to prepausal obstruents, which become unvoiced as no chord vibration can be assimilated from a following sound. Final devoicing and voicedness assimilation are therefore two aspects of the same regressive spreading movement. The underlying voicedness of final obstruents can be determined only by checking alternating forms:

prepausal position	alternating with	underlying form	followed by unvoiced consonant	followed by voiced consonant
<i>cos</i> ['kɔs] 'body'	<i>cossos</i> ['kosus] 'bodies'	/kɔs/	<i>cos fort</i> [kɔs'fɔrt] 'strong body'	<i>cos dèbil</i> [kɔz'ðɛβit] 'weak body'
<i>cas</i> ['kas] 'case'	<i>casos</i> ['kazu] 'cases'	/kaz/	<i>cas pràctic</i> [kas'praktik] 'practical case'	<i>cas nou</i> [kaz'nɔw] 'new case'
<i>mut</i> ['mut] 'speechless (masc. sg.)'	<i>muda</i> ['muðə] 'speechless (fem. sg.)'	/mud/	... <i>mut té</i> ... [mut'te] '...speechless has...'	<i>mut de</i> ... ['muddə] 'speechless of...'
<i>tot</i> ['tot] 'all (masc. sg.)'	<i>tot</i> ['tota] 'all (masc. sg.)'	/tot/	<i>tot trencat</i> [ˌtɒttrən'kat] 'all broken'	<i>tot dret</i> [tɒt'drɛt] 'all straight'

Despite the common behaviour in cases like the ones above, stops and fricatives differ when a following word or a lexical morpheme of a given set begins with a vowel. In such cases, stops become unvoiced, but fricatives become voiced: *cos inflat* [kəziŋ 'flat] 'swollen body', *el mateix any* [əlməte'ʒaŋ] 'the same year' (cf. *mateixos* [mə 'tefus]), but *fred intens* [fretɪŋ'tens] 'intense cold', *nord-americà* [nɔrtəməri'ka] 'North American'. The voicing rule applying to prevocalic word-final fricatives seems not to apply so commonly to final /f/ (Recasens i Vives 2014, 253).

Another important word-final process affecting obstruents deals with /v/: when in final or pre-final position (i.e. followed by a final /+s/) it turns into the velar glide ([w]): *nova* ['nɔvə] (or ['nɔβə]: see 1.2 above on the alternation [v]/[b] (or [β])) 'new' (fem. sg.) / *nou* ['nɔw] 'new' (masc. sg.), *nous* ['nɔws] 'new' (masc. pl.).

Assimilation is general when it affects the point of articulation of the alveolar nasal. For instance, the /n/ in *un* (indefinite article) can accommodate its point of articulation to the following consonant: *un senyor* ([n]) 'a gentleman', *un bou* ([m]) 'an ox', *un falcó* ([ŋ]) 'a falcon', *un tauró* ([ŋ]) 'a shark', *un xai* ([ŋ]) 'a lamb', *un gos* ([ŋ]) 'a dog'. The bilabial nasal undergoes the same adjustments in Majorcan and Minorcan Catalan, although it maintains its point of articulation in other dialects. In all dialects /ɫ/ is pronounced as dental ([ʎ]) before /t/ or /d/. In Majorcan and Minorcan coda stops completely assimilate to a following consonant, except if this is a sibilant. In such a case, the stop and the sibilant merge into an affricate maintaining the point of articulation of the sibilant. The following are examples of assimilation and affricate formation: *tub gros* [tug'grɔs] 'big tube', *tub nou* [tun'nɔw] 'new tube', *tub lleuger* [tuʎʎəw'ze] 'light tube', *tub girat* [tuðʒi'rat] 'turned tube'. This tendency is shared with Northern Catalan, except for velars, which retain their point of articulation. In careless speech, alveolar stops (/t/, /d/) also undergo assimilation to a following consonant in all dialects, except in those where cluster simplification precludes the contact, as in Valencian and Alguerese.

Dissimilation affects alveolar sibilants in Majorcan and Minorcan whenever they face a following sibilant (either alveolar or prepalatal). In such cases, the first turns into a stop and, consequently, an affricate arises: *dues senyores* [duəts:ə'ɲorəs] 'two ladies', *dos gelats* [dodʒ:ə'ɫats] 'two ice creams', *tres xinesos* [trətʃi'nəzɔs] 'three Chinese'. The form of the Majorcan and Minorcan definite article for the masculine plural when preceding a word-initial vowel can be explained on these grounds (Dols 2002), as the lexical /s/ of Balearic article (*es/sa/es/ses*) collides with an inflectional /+s/ (a plural marker). This results, for example, in *es avions* (underlyingly /s+s#əviɔn+s/) surfacing as [ədʒ:əvi'ɔns] where the application of three rules is visible: dissimilation of sibilants, voicedness assimilation and the strengthening of affricates (see below for the latter).

The weakening process known as "spirantisation", which affects Portuguese and Spanish to varying degrees, also occurs in Catalan where voiced stops become approximants in a context of continuancy, i.e. when during pronunciation the airflow of sounds surrounding the stop is intense or the articulators are wide open. This is usually

the case with vowels and glides, fricatives (except when a fricative follows and then an affricate arises), rhotics, and laterals (except when a lateral precedes /d/): *cada dia* [ˌkaðəˈðia] ‘every day’, *agregar* [əɣrəˈɣa] ‘to aggregate’, *albada* [əlˈβaðə] ‘dawn’, but *adjectiu* [ədʒːəkˈtiw] ‘adjective’, *el dia* [əlˈdiə] ‘the day’, *anguila* [ənˈɡiɫə] ‘eel’. Spirantisation is a common phenomenon in Catalan, although the exact degree of adherence to the rule may depend on the specific dialect: /b/ → [β] seems not to be general in the areas where [v] exists, Alguerese may lack /g/ → [ɣ] according to Recasens i Vives (1991, 234–235), and a further weakening of the intervocalic /d/ is observed in the same dialect, with /d/ becoming [r] or [ɹ]. This process goes even further in Southern Valencian, which exhibits elision of intervocalic /d/, especially in past participles and, in a narrower area, in other contexts (Colomina i Castanyer 1985, 109–127).

The intervocalic [j] can undergo weakening in Majorcan Catalan, ranging from centralisation to total elision, depending on the precise region. In this case, elision is the most common solution in Minorcan Catalan: *palla* (expected [ˈpaɲə], see 1.2 above on this spelling) [ˈpaɛə] / [ˈpaə] ‘straw’, *vella* [ˈveə] ‘old’ (fem. sg.) (cf. *vell* [ˈvej] ‘old’ (masc. sg.)).

4 Morphophonemics

4.1 Elision

The elision of final /-r/ applies to verbs (infinitives) and nouns, which maintain the underlying rhotic, as can be seen when a clitic is added to the right edge of an infinitive or a morpheme beginning with a vowel is added to the end of a noun: *anar* [əˈna] ‘to go’/*anar-hi* [əˈnari] ‘to go there’, *dir* [ˈdir] ‘to say’/*dir-m’ho* [ˈdirmu] ‘to say it to me’, *escriptor* [əskripˈto] ‘writer’ (masc. sg.)/*escriptora* [əskripˈtora] ‘writer’ (fem. sg.). Dialects adhere to this rule to varying degrees; it is largely followed in Majorcan Catalan, where only a few monosyllabic words or non-traditional ones avoid it (as *pur* ‘pure’, *per* ‘for’, *militar* ‘military’, *motor* ‘engine’; but *mar* [ˈma] ‘sea’, *cor* ‘heart’ [ˈkɔ], *or* [ˈɔ] ‘gold’), and it is generally not applied in Valencian (exceptions in the most Northern and Southern parts, Colon 1952; Colomina i Castanyer 1985). In the other regions where it is applied, final /r/ persists in monosyllables, some abstract words ending in *-or*, and adjectives derived by means of suffixes: *-ar* (*familiar*, *popular*, etc.), *-er* (only learned words: *auster*, *sever*), *-ior* (*anterior*, *superior*), or in the verbal forms *mor* ‘s/he dies’, *mors* ‘you die’ (Bibiloni 2002, 278). Wherever the final *-r* is elided in Western dialects, it does not reappear in verbs before a following clitic, which takes the normal form for postvocalic clitics: *matar-lo* [maˈtal] ‘to kill him’, *donar-lo* [doˈnal] ‘to give it’. The same can be said for paroxytones in Central Catalan: *conèixer-nos* [kuˈnɛʃəns] ‘to know each other’ (1st pers. pl.), *córrer-hi* [ˈkorəj]/[ˈkɔri] ‘to run there’.

Many oxytone nouns ending in a vowel maintain an underlying /n/ and elide it when final, but it emerges when a morpheme is added due to inflection or derivation:

mà/mans ('hand'/'hands'), *bo/bona/bons/bones* ('good' masc. sg./fem. sg./masc. pl./fem. pl.), *raó* ('reason')/*raonable* ('reasonable'). If a shortcut rule is to be stated such as 'add an [n] for forms relating to a vowel final oxytone', then some caution must be taken: except for *bé* 'well' (cf. *benestar* 'welfare'), and colloquial forms relating to *així* 'so' (cf. *aixins* 'so' (dialectal, colloquial)), *aquí/allà* (cf. *just su aquines/allanes* 'exactly here/there' (dialectal, colloquial)) adverbs do not exhibit this behaviour when they receive a morpheme, even if they play the role of nouns: *sís* 'yes' (pl.), *demàs* 'tomorrow'. An exception to this rule of thumb would be items not deriving from a Latin word with an /n/ at the end of the stem: *fe* 'faith', *mercè* 'mercy', *bisturí* 'lancet', *cafè* 'coffee', *sofà* 'sofa', *nu* 'undressed', *cru* 'raw', etc. In Western dialects, Camp de Tarragona, Eivissenc and Alguerese, the /-n/ can reappear in the plural of certain paroxytones (*hòmens* 'men', *jòvens* 'youngsters', *òrfens* 'orphans', *còvens* 'basket'), whereas in the other dialects it only reappears in words derived through suffixation (*homenàs* 'great man', *jovenesa* 'youth', *orfenesa* 'orphanhood', *covenet* 'basket' (diminutive)). Proparoxytones, learned (par)oxytones and certain proclitic adjectives or determiners appear resistant to the elision rule: *bàdminton* 'badminton', *èpsilon* 'epsilon', *anglòfon* 'Anglophone', *origen* 'origin', *abdomen* 'abdomen', *edèn* 'Eden', *un* 'a', *algun* 'certain', *mitjan* 'middle', *quin* 'which', etc.

4.2 Epenthesis

Masculine nouns ending in /s/ or /z/ (written <s> or <ç>) insert an [o] before the plural /+s/, as in *gas* 'gas' – *gasos* 'gases', *nas* 'nose' – *nassos* 'noses' (see 3.2.2 above on voicedness in sibilants), *fal·laç* 'mendacious' (masc. sg.) – *fal·laços* 'mendacious' (masc. pl.). The same can be said about those ending in /ʃ/ (<ix>), except for the (rare) option of adding a simple [s] with no epenthetic vowel: *mateix* 'same' (masc. sg.) – *mateixos* / *mateixs* 'same' (masc. pl.). The unvoiced palatal affricate (written <tx>) always requires an epenthetic vowel *despatx* 'office' – *despatxos* 'offices'. However, when it is derived from a voiced palatal sibilant (graphically rendered as <(i)g>, see 3.2.2 on this) the insertion of [o] is optional. When the [o] is inserted, a fricative or an affricate is produced, depending on the underlying contrast erased at the word edge following the affrication and devoicing rules explained in 3.2.2: *lleig* 'ugly' (masc. sg.) – *lleigs/lletjos* [ˈlɛdʒus] 'ugly' (masc. pl.), *passeig* 'walk' – *passeigs/passejos* [pəˈsɛʒus] 'walks'. Masculine noun endings [sp], [st], [sk], [kst] (<sp>, <st>, <sc>, <xt>) permit both simple addition of plural /+s/, and insertion of [o] before it, as in *creps/crepos* 'rippled' (masc. pl.) *llests/llestos* 'ready' (masc. pl.), *toscs/toscós* 'rough' (masc. pl.), *mixts/mixtos* 'mixed' (masc. pl.). *Aquest* only allows simple addition of /+s/, although [əˈkɛstus] it is not unknown colloquially. As stated above, this type of epenthesis occurs only in masculine forms. For this reason, the plural of nouns such as *post* 'board', *forest* 'forest' or *falç* 'sickle', all of them feminine, do not allow an epenthetic vowel: *posts*, *forests*, *falçs*. The same can be said for a short set of masculine parox-

ytones [kst] (<x>): *índex/índexs* ‘index/indices’, *apèndix/apèndixs* ‘appendix/ appendices’. Adjectives sharing the same form ending in /s/ (<ç>) in the singular exhibit two forms in the plural according to gender: *atroç* ‘atrocious’, *audaç* ‘bold’, *capaç* ‘able’, *eficaç* ‘effective’, *feliç* ‘happy’, etc. end in -ços ([sus]) or -ces ([səs]) (Dols/Mansell 2017, 22–24).

4.3 Phonology in verb paradigms

A verbal ending added to a root can create an ill-formed syllable structure. In such cases, epenthesis can solve the problem, as in *coneixes* [ku'neʃəs] ‘you know’ (sg.), *vences* [vensəs] ‘you win’ (sg.) (cf. *reps* [rɛps] ‘you receive’ (sg.)), *cuses* [kuzəs] ‘you sew’ (cf. *escups* [əs'kups] ‘you spit’), *corre* [korə] ‘she/he/it runs’.

As in etymological derivation CINERE > [sɛnrə] (still active in Majorcan Catalan with a vibrant, [sɛnrə]) > *cendra* [sɛndrə], [d] is inserted in synchronic conjugation to avoid contact between /ʎ/ and /r/, and /n/ and /r/: *moldre* ‘to grind’, *moldré* ‘I shall grind’, *moldria* ‘I/she/he/it would grind’ (cf. *molen* ‘they grind’, *molíem* ‘we ground’ (imperfect)), *vendre* ‘to sell’, *vendré* ‘I shall sell’, *vendria* ‘I/she/he/it would sell’ (cf. *venen* ‘they sell’, *veníem* ‘we sold’ (imperfect)).

An important issue regarding the morphology of verbs arises in Balearic and Alguerese Catalan. Much like in the subjunctive, where no mark for 1st pers. sg. is added (e.g. *canti* ‘I sing’ (subj.)), where /i/ is the mark for the subjunctive mood), 1st pers. sg. in the indicative is pronounced with no mark for person in the 1st and partially in the 2nd and 3rd conjugations, similar to the 2nd and the 3rd conjugation in Valencian (*rep* ‘I receive’, *dorm* ‘I sleep’ vs. Central Catalan *rebo*, *dormo*). There is no problem when the syllabic structure adjusts to the predictions of the sonority scale (see 2.2 above): *cant* ‘I sing’, *torn* ‘I return’, *parl* ‘I speak’. When the sonority scale is challenged, unusual structures arise, and they are treated in different ways. In Balearic the contrast between rhotics is maintained at the right end of the word, two identical consonants are simplified to one, and disyllabic clusters are pronounced only before a pause and reduced when preceding a word-initial consonant: *emparr* [əm'par] ‘I arrange plants in an arbour’/ *empar* ‘I protect’, *vacil* ‘I hesitate’, cf. *vacil·lar* ‘to hesitate’, but *vetlar* [vət'la] ‘to watch’ > *vetl* [vət'] ‘I watch’ (Bibiloni 1983, 121)), *entr* [əntɾ] ‘I go in’, *infl* [im'fɫ] ‘I inflate’ (voicelessness of final sonorants according to Recasens i Vives (1991, 310 and 327)), but *entr molt* [əm'moɫt] ‘I go in often’, *infl poc* [im'pɔk] ‘I inflate little’. When the verb stem ends in a vowel, a glide is added. The glide is chosen according to the quality of the preceding vowel: rounded vowels select /w/ and unrounded select /j/, as in *crei* [krej] ‘I create’, *estudii* [əstu'ðij] ‘I study’/ *llou* [ləw] ‘I praise’, *suu* [suw] ‘I sweat’ (see Dols/Wheeler 1996 for a study on the syllable structure in these cases).

5 Suprasegmentals

5.1 Stress assignment

At first glance, Catalan seems to have lexical stress, that is to say that one is forced to learn the place of stress in each individual word. However, such an assumption does not seem to take evident regularities into account. As Serra (1996) and Oliva/Serra (2002) point out, the typical stress patterns in Catalan are as stated in (10):

(10a) Words ending in a consonant are oxytones: *estret* 'narrow', *remolc* 'trailer', *arrap* 'scratch', *velam* 'sails', *contraban* 'smuggling', *viarany* 'path', *estel* 'kite', *conill* 'rabbit', *babau* 'stupid', *espai* 'room' all carry the stress on the last syllable.

(10b) Words ending in a vowel are paroxytones: *casa* 'house', *colze* 'elbow', *cossi* 'tub', *carro* 'cart', *tribu* 'tribe' are all stressed on the penultimate vowel.

However, these patterns do not cover the entire inventory of stressed words, as exceptions to them seem to exist, as in (11):

(11a) Oxytones ending in a vowel: *català* 'Catalan', *puré* 'mash', *alè* 'breath', *alpí* 'alpine', *turó* 'hill', *ressò* 'echo', *cadascú* 'everyone', etc.

(11b) Paroxytones ending in a consonant: *àpat* 'meal', *càndid* 'candid', *divendres* 'Friday', *aborigen* 'aborigin', *pètal* 'petal', *fotògraf* 'photographer', *màgic* 'magic (adj.)', etc.

(11c) Proparoxytones exist: *ingènua* 'naive (fem. sg.)', *llàstima* 'pity', *èxode* 'exodus', etc.

Apart from these cases, verbal inflection shows some peculiarities, as in 1st and 3rd pers. sg. future: *menjaré* 'I shall eat', *menjarà* 'she/he/it will eat', with a final stressed vowel, or *dormies*, *dormíem*, *dormíeu*, *dormien*, *dormires*, *dormírem*, *dormíreu*, *dormíssi*, *dormíssis*, *dormíssim*, *dormíssi*, *dormíssin* 'to sleep' (different persons and tenses) where the penultimate vowel is stressed and all forms end with a consonant (recall that the <u> in <-eu>, <-iu> is a glide).

A number of apparent exceptions in (11a) can be discarded on the basis that stress is in fact assigned to morphemes, not to words. According to this, all cases of stressed final vowels in words with an underlying final /-n/ visible in alternations (*català*/*catalana*, see 4.1 on elision) are not exceptional. Other words, as *puré*, *bisturí*, *cafè*, *sofà*, etc. are true exceptions to the general stress pattern.

Some cases in (11b) are known as "preaccented" suffixes, i.e. suffixes requiring that stress be assigned to the immediately preceding syllable, such as *-fon*, *-graf*, *-ic*, *-fil*, *-fob* (Mascaró 1986): *tele* + *graf* → *telègraf* + *ic* → *telegràfic*; *anglo* + *fon* → *anglòfon*. Not only do these suffixes require stress on the preceding syllable, but there is also an observable preference for open vowels in that syllable; there is no reason to propose underlying open or mid-open vowels if any close-mid or mid vowel will be open in such a situation. Other cases such as *pètal*, *divendres*, etc. are lexically marked as

having an extrametrical final consonant: *divendre(s)*. When stress is assigned, the final syllable is light and therefore the penultimate syllable receives it.

Derivational and inflectional morphemes can add complexity to stress assignment; an inflectional morpheme such as feminine /+ə/, which is not lexically marked for stress and does not end in a consonant, can be added to a lexical morpheme such as *ingenu*, which regularly has stress placed on its penultimate syllable, to form the feminine singular form of the adjective: *ingènua* [iŋ'ʒenuə]. In cases such as *ingènua* – with a high vowel at the end of the stem –, and even more commonly when the last consonant is a sibilant, the pressure to turn the high vowel into a glide is stronger (see 1.2 and 2.1 above on i/j and u/w alternations): *ingènua* [iŋ.'ʒɛ.nwə] 'naive', *gràcia* ['gra.sjə]. Other cases of proparoxytones, such as *èxode*, *hàbitat*, etc. are considered to have a final extrametrical syllable, not counted in the stress-assignment process: *èxo (de)*, *hàbi(tat)*.

When derivational suffixes already stressed are added to the right edge of a stem, the rightmost stress becomes the word stress, as in /'bɔn+'dat/ → [buŋ'dat], /'kati(d)+'ɛz+ə/ → [kəli'ðɛzə]. Compounds can hold two stresses, as can adverbs of manner ending in *-ment*, with the result that the stressed vowel in the leftmost root is protected against vowel reduction (see above 3.2.1): *lentament* [lɛŋtə'men] (*[lɛŋtə'men]) 'slowly', *camallarg* [kamə'lark] (*[kəmə'lark]) 'long-legged'.

5.2 Intonation

Tone not being of phonological importance in Catalan, the functions of intonation are mainly syntactic and pragmatic. Declarative sentences usually finish with falling intonation after rising at the end of the subject, as in Figure 2, the intonation contour for *La Maria menja sopa* 'Maria is eating soup'

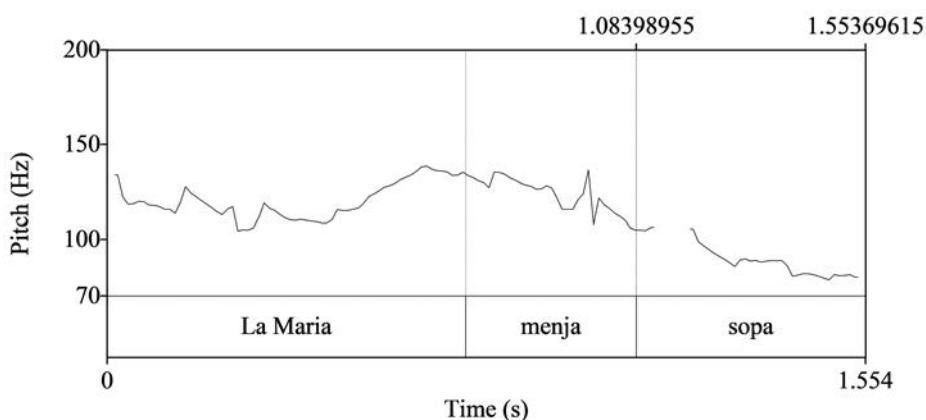


Figure 2: Intonation contour for *La Maria menja sopa* 'Maria is eating soup'.

Observe how after *La Maria*, the pitch decreases until it reaches its lowest level, which indicates the end of the sentence. If the sentence is lengthened by means of a new phrase, then the whole contour is rebuilt, and where we first found a low final pitch, we now find a much higher one, higher than that of the subject just to indicate that the sentence is going to be completed, as in Figure 3 for *La Maria menja sopa i bistec amb patates* 'Maria is eating soup and steak and chips'

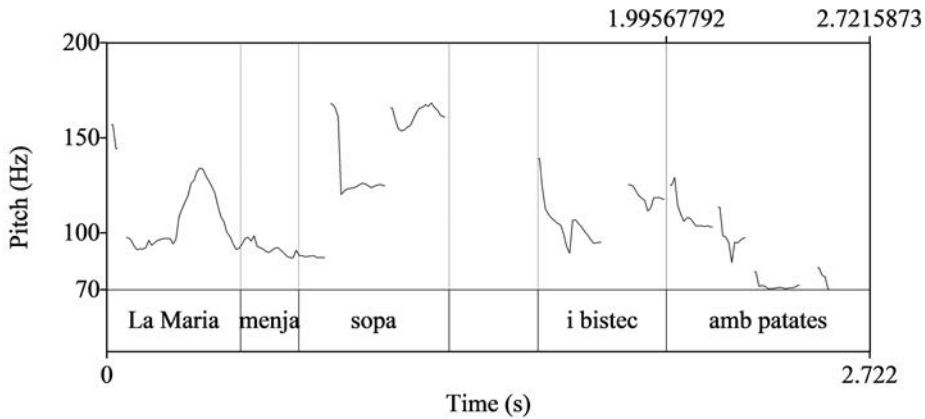


Figure 3: Intonation contour for *La Maria menja sopa i bistec amb patates* 'Maria is eating soup and steak and chips'.

Somewhere between syntax and pragmatics, focused information is marked by position as well as pitch; an element focused on the left boundary of the sentence appears in a shape similar to that of subjects, i.e. with a rising pitch-curve, such as in the example in Figure 4, *El llibre, no l'ha llegit* 'The book, s/he hasn't read it'.

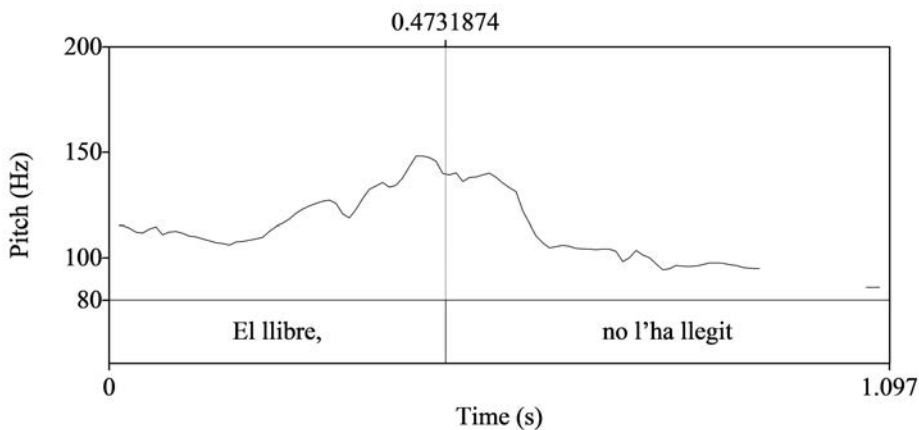


Figure 4: Intonation contour for *El llibre, no l'ha llegit* 'The book, s/he hasn't read it'.

Elements focused at the right edge exhibit a low pitch, maintained up to the end of the sentence, as in Figure 5 for *No ho sé, quan ha tornat* ‘I don’t know what time s/he came back’

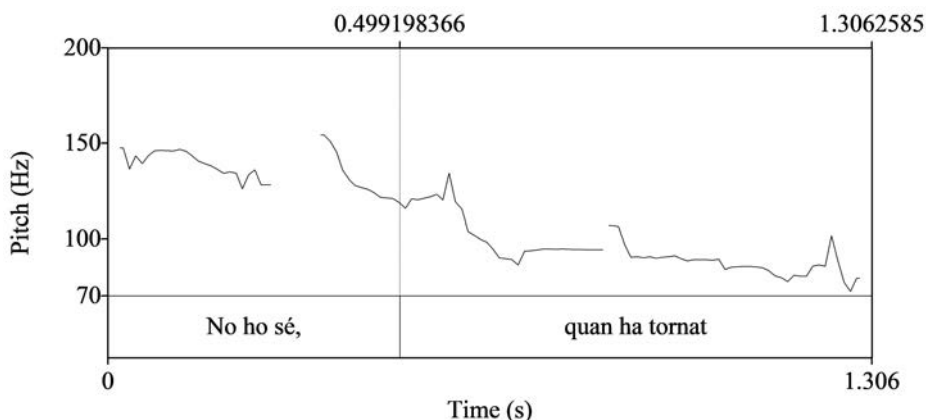


Figure 5: Intonation contour for *No ho sé, quan ha tornat* ‘I don’t know what time s/he came back’.

Intonation in questions is open to dialectal variation. In polar questions both rising and falling contours exist (Prieto/Cabré 2007–2012; 2013) depending on the dialect. Both contours appear in Central and Northwestern Catalan, again depending on the specific region. Valencian and Northern Catalan stick to the rising contour (Figure 6a), while Balearic and Alguerese prefer falling polar questions (Figure 6b): *Has vist la Maria?* ‘Have you seen Maria?’

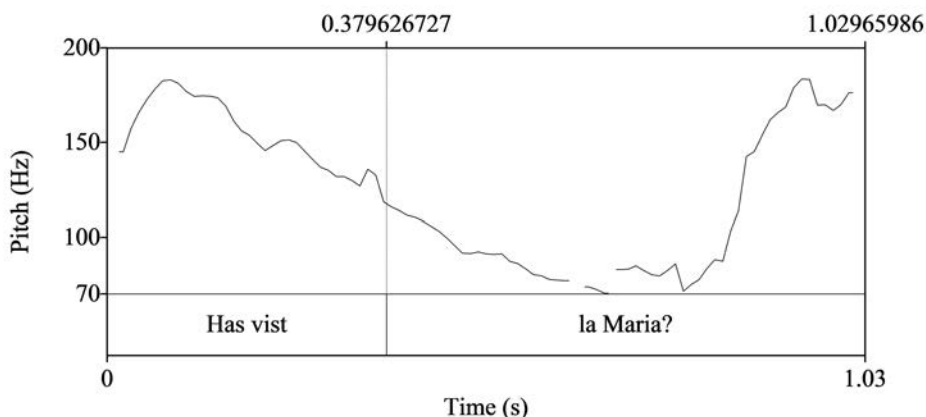


Figure 6a: Intonation contour for *Has vist la Maria?* ‘Have you seen Maria?’.

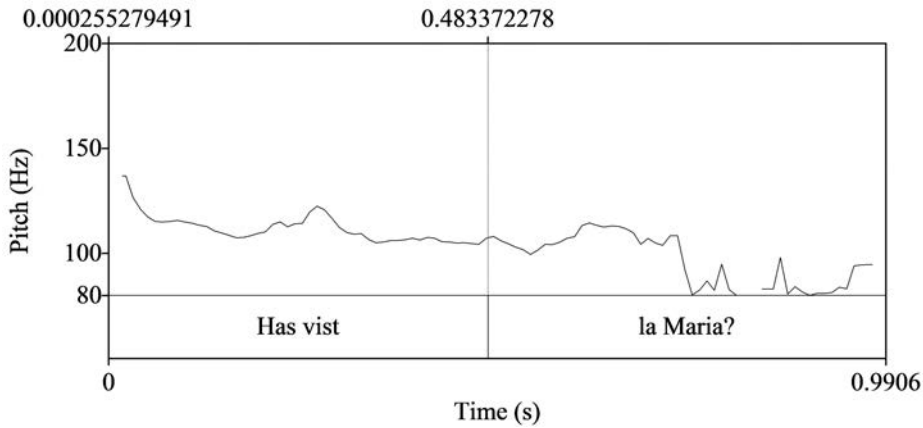


Figure 6b: A different intonation contour for *Has vist la Maria?* 'Have you seen Maria?'.

Non-polar questions display a falling end, after a high tone maintained up until the nuclear syllable. This syllable can be a part of the maintained high tone or can belong to the falling contour depending on the dialect (Figure 7).

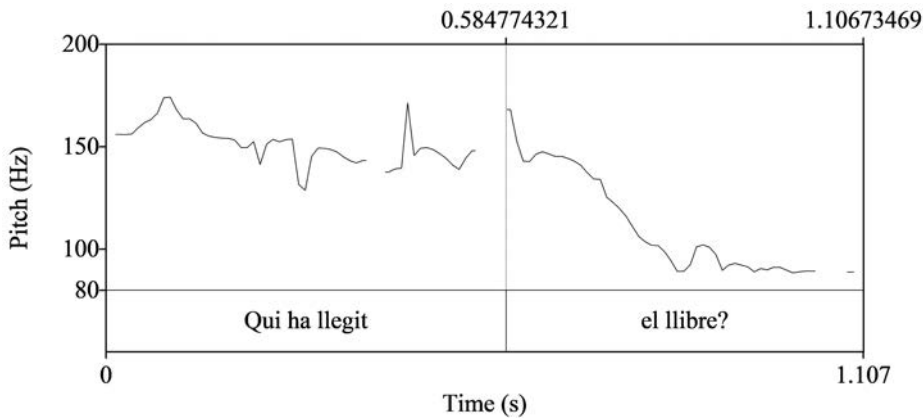


Figure 7: Intonation contour for *Qui ha llegit el llibre?* 'Who has read the book?'.

In contrast with declarative sentences, exclamation normally exhibits a rising pitch at the end, together with a higher prominence of other peaks in the sentence, as in Figure 8, where the final rising pitch and the prominence of the stressed syllable in *llibre* should be pointed out.

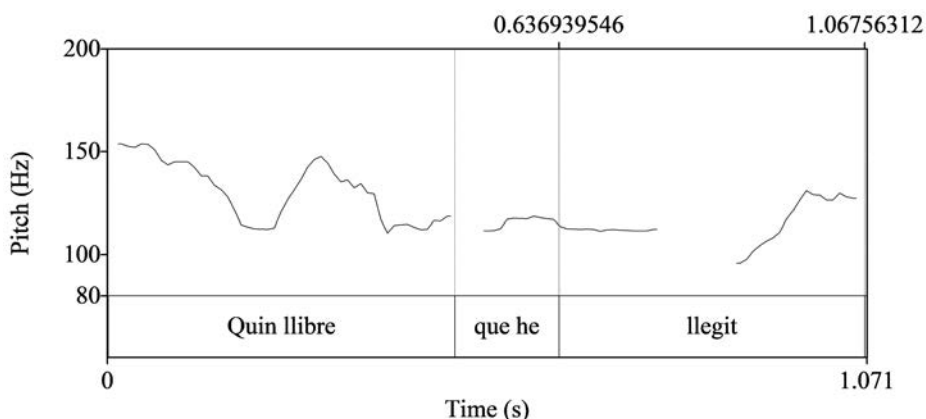


Figure 8: Intonation contour for *Quin llibre que he llegit* 'What a book that I have read'.

6 Bibliography

- Alcover, Antoni M. (1908), *Una mica de dialectologia catalana*, Bolletí del Diccionari de la Llengua Catalana 1, 209–560.
- Badia i Margarit, Antoni M. (1962), *Gramàtica catalana*, Madrid, Gredos.
- Bargalló Valls, Josep (2007), *Què és la mètrica*, Barcelona, Edicions 62.
- Bibiloni, Gabriel (1983), *La llengua dels mallorquins. Anàlisi sociolingüística*, Barcelona, Universitat de Barcelona.
- Bibiloni, Gabriel (2002), *Elisió de -n i -r, distribució de les ròtiques i altres fenòmens consonàntics en el mot*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 1, Barcelona, Empúries, 271–285.
- Bibiloni, Gabriel (2016), *El català de Mallorca. La fonètica*, Palma, Muntaner.
- Bonet, Eulàlia/Lloret, Maria-Rosa (1998), *Fonologia catalana*, Barcelona, Ariel.
- Cardona, Osvald (1977), *Els grups de vocals en contacte*, Barcelona, Fundació Vives Casajuana.
- Colomina i Castanyer, Jordi (1985), *L'alcantí. Un estudi sobre la variació lingüística*, Alacant, Institut d'Estudis Joan Gil-Albert.
- Colon, Germà (1952), *Unes notes sobre la pèrdua de la -R final etimològica*, Revista Valenciana de Filologia 2/1, 57–65.
- Coromines, Joan (1971), *Lleures i converses d'un filòleg*, Barcelona, Club Editor.
- Dols, Nicolau (2000), *Teoria fonològica i sil·labificació. El cas del català de Mallorca*, Palma, Universitat de les Illes Balears.
- Dols, Nicolau (2002), *Fenòmens en grups consonàntics*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 1, Barcelona, Empúries, 319–343.
- Dols, Nicolau (2011), *Calculant distàncies. Els límits de l'harmonia i de la coarticulació CVC en el català de Mallorca*, in: Maria-Rosa Lloret/Clàudia Pons (edd.), *Noves aproximacions a la fonologia i la morfologia del català. Volum d'homenatge a Max W. Wheeler*, Alacant, Institut Interuniversitari de Filologia Valenciana, 175–196.
- Dols, Nicolau/Mansell, Richard (2017), *Catalan. An Essential Grammar*, Oxon/New York, Routledge.
- Dols, Nicolau/Wheeler, Max W. (1996), *El consonantisme final del mallorquí i el "llicenciament d'obertures"*, Caplletra 19, 51–64.

- Fabra, Pompeu (1891), *Ensayo de gramática de catalán moderno*, Barcelona, Est. y Lib. L'Avenç de Massó y Casas.
- GIEC = Institut d'Estudis Catalans (2016), *Gramàtica de la llengua catalana*, Barcelona, Institut d'Estudis Catalans.
- Jakobson, Roman/Halle, Morris (1956), *Fundamentals of Language*, The Hague, Mouton & Co.
- Jiménez, Jesús (1998), *Valencian Vowel Harmony*, *Rivista di Linguistica* 10, 137–161.
- Julià, Joan (1981), *Estudi contrastiu dels oclusius de l'anglès i del català. Un experiment acústic*, *Estudi General* 1/2, 75–85.
- Kiparsky, Paul (1979), *Metrical Structure Assignment is Cyclic*, *Linguistic Inquiry* 10, 421–442.
- Lleó, Conxita (1970), *Problems of Catalan Phonology*, Seattle, University of Washington.
- Mascaró, Joan (1986), *Morfologia*, Barcelona, Enciclopèdia Catalana.
- Mascaró, Joan (1987), *Syllable-Final Processes in Catalan*, in: Carol Neidle/Rafael A. Núñez-Cedeño (edd.), *Studies in Romance Languages*, Dordrecht/Providence, Foris, 133–146.
- Mascaró, Joan (1991), *La importància lingüística de la polèmica sobre la vocal neutra en català central i el seu caràcter fonemàtic*, *Els Marges* 44, 33–43.
- Mascaró, Joan (2002), *El sistema vocàlic. Reducció Vocàlica*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 1, Barcelona, Empúries, 89–123.
- Montoya, Brauli (1981), *Confluència de llengües a les valls del Vinalopó (un cas de sociolingüística valenciana)*, Alacant, Universitat d'Alacant.
- Oliva, Salvador (1980), *Mètrica catalana*, Barcelona, Quaderns Crema.
- Oliva, Salvador (1988), *Introducció a la mètrica*, Barcelona, Quaderns Crema.
- Oliva, Salvador (1992), *La mètrica i el ritme de la prosa*, Barcelona, Quaderns Crema.
- Oliva, Salvador/Serra, Pep (2002), *Accent*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 1, Barcelona, Empúries, 345–391.
- Palmada, Blanca (1994), *La fonologia del català. Els principis generals i la variació*, Barcelona, Universitat Autònoma de Barcelona.
- Pons Moll, Clàudia (2007), *La teoria de l'optimitat. Una introducció aplicada al català de les Illes Balears*, Barcelona, Institut Menorquí d'Estudis/Publicacions de l'Abadia de Montserrat.
- Pradilla Cardona, Miquel-Àngel (2002), *Ensordiment, espirantització i fenòmens que afecten les sibilants*, in: Joan Solà et al. (edd.) *Gramàtica del català contemporani*, vol. 1, Barcelona, Empúries, 287–318.
- Prieto, Pilar/Cabré, Teresa (edd.) (2007–2012), *Atles interactiu de l'entonació del català*, <http://prosodia.upf.edu/atlesentonacio/> (last accessed: 13.02.2018).
- Prieto, Pilar/Cabré, Teresa (edd.) (2013), *L'entonació dels dialectes catalans*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Prince, Alan/Smolensky, Paul (1993), *Optimality Theory. Constraint Interaction in Generative Grammar*, University of Rutgers/University of Colorado.
- Recasens i Vives, Daniel (1990), *Temes de fonètica i lingüística del català*, Tarragona, Publicacions de la Diputació.
- Recasens i Vives, Daniel (1991), *Fonètica descriptiva del català*, Barcelona, Institut d'Estudis Catalans.
- Recasens i Vives, Daniel (1993), *Fonètica i fonologia*, Barcelona, Enciclopèdia Catalana.
- Recasens i Vives, Daniel (2014), *Fonètica i fonologia experimentals del català. Vocals i consonants*, Barcelona, Institut d'Estudis Catalans.
- Roca, Iggy/Johnson, Wyn (1999), *A Course in Phonology*, Oxford, Blackwell.
- Selkirk, Elisabeth O. (1982), *Syllables*, in: Harry van der Hulst/Norval Smith (edd.), *The Structure of Phonological Representations*, Dordrecht, Foris, 337–383.
- Serra, Pep (1996), *La fonologia prosòdica del català*, Girona, Universitat de Girona.
- SIL, “diphthong”, *Glossary of Linguistic Terms*, <http://www.glossary.sil.org/term/diphthong> (last accessed: 13.02.2018).

- Vallverdú Albornà, Teresa (2002), *Fenòmens en grups vocàlics*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 1, Barcelona, Empúries, 125–167.
- Vennemann, Theo (1988), *Preference Laws for Syllable Structure*, Berlin, Mouton de Gruyter.
- Wheeler, Max W. (1979), *Phonology of Catalan*, Oxford, Blackwell.
- Wheeler, Max W. (1987), *L'estructura fonològica de la síl·laba i el mot en català*, in: *Miscel·lània Antoni M. Badia i Margarit*, vol. 6, Barcelona, Publicacions de l'Abadia de Montserrat, 79–108.
- Wheeler, Max W. (2005), *The Phonology of Catalan*, Oxford, Oxford University Press.

5 Morphosyntax

Mar Massanell i Messalles

5.1 Word Classes, Inflectional Categories and Paradigms

Abstract: This chapter starts out by defining the basic units of morphology, namely *word* and *morpheme*, and explaining the distinctions between *root*, *derivational affix* and *inflectional marker*, together with the concepts of *stem* and *ending*. In the second section, we deal with word classes. First, we describe the nine main word classes found in Catalan – nouns, adjectives, determiners, verbs, adverbs, pronouns, prepositions, conjunctions and interjections – and some transversal word classes – demonstratives, possessives, quantifiers, relatives, interrogatives, exclamationatives and connectors. Then we introduce the dichotomy between closed and open classes, and between variable and invariable classes. In the third section, we present the categories of nominal inflection – gender, number, case and person – and verbal inflection – conjugation, tense, aspect, mood, person and number – found in Catalan. In the fourth section, we deal with paradigms and the markedness oppositions that may be found within them, and also some phenomena which can affect paradigms such as defectiveness and syncretism. Section five introduces some concepts related to the phonological realisation of morphemes, and specifically the concepts of *morph*, *zero morph*, *allomorphy* and *fusion*. At this point, we introduce the Catalan inflectional markers in nominal word classes – that is, gender and number markers – and then list the closed classes of articles, demonstratives, possessives and personal pronouns. Finally, we present the inflectional markers that signal conjugation, tense and person in Catalan verbs.

Keywords: word classes, inflectional markers, nominal inflection, verbal inflection, paradigms

1 Words and morphemes

There are two basic units of morphology: words, like *tauleta* ‘small table’, and morphemes, like the constituents *taul-*, *-et-* and *-a*. Words are the units of meaning of a language that can act with syntactic autonomy to form phrases and sentences, as in *tauleta blanca* ‘small white table’ and *M’agrada aquella tauleta blanca* ‘I like that small white table’. Thus, words are simultaneously the largest units of morphology and the smallest units of syntax. Meanwhile, morphemes are the smallest units of

<https://doi.org/10.1515/9783110450408-006>

language with meaning or grammatical function, as in *taul-* ‘table’, *-et-* ‘small’, *-a* [+feminine]. Usually, words consist of several morphemes grouped together (*taul+et+a*), although sometimes a word consists of a single morpheme, as in *gos* ‘dog’, *dorm* ‘(he/she/it) sleeps’ or *sota* ‘under’, which all appear in the sentence *El gos dorm sota la tauleta blanca* ‘The dog is sleeping under the small white table’.

1.1 Roots, affixes and morphological processes

Polymorphic words consist of a basic morpheme – the one which contributes most to the lexical meaning of the word – which is called the *root*, and then one or more attached morphemes which add different sorts of semantic content or grammatical specifications. These attached morphemes are called *affixes* and can be either derivational or inflectional. Derivational affixes serve to form new words; for example, *-et-* provides the word *taula* ‘table’ with diminutive value, to yield *tauleta* ‘small table’. The derivational affixes of Catalan will be presented in detail in another chapter (§7.2 Word-Formation). By contrast, inflectional affixes are markers of grammatical properties, like feminine gender or plural number, which is why they are called *markers*, as in *feminine marker*, *plural marker*, etc. These inflectional markers will be the focus of the next few pages.

In Catalan, inflectional affixes are always suffixes, that is, they are placed at the end of the word. The set of markers added to the end of a word is referred to as its *ending*. Since several morphological processes can be applied successively, we use the term *stem* to refer to the word constituent to which inflectional processes are applied. If the stem coincides with the root, it is called *monomorphemic*, as in simple words like *taula* ‘table’. On the other hand, the stem is called *polymorphemic* if it is made up of the root and one or more derivational affixes, as in derived words like *tauletes*, whose morphological segmentation is illustrated below.

$$[[\text{taul}_{\text{root ('table')}} + \text{et}_{\text{derivational affix ('small')}}]_{\text{stem}} + [\text{e}_{\text{feminine marker}} + \text{s}_{\text{plural marker}}]_{\text{ending}}]_{\text{inflected word}}$$

The stem is also regarded as polymorphemic if it consists of more than one root, as is the case in compounds such as *picaporta* ‘door-knocker’, for example.

2 Word classes

From the viewpoint of morphology, words can be classed either as inflective, or *variable*, if they can take inflectional affixes, like *nen* ‘boy’, *nen+s* ‘boys’, *nen+a* ‘girl’, *nen+e+s* ‘girls’, or as non-inflective, or *invariable*, if they cannot take inflectional markers, like *allà* ‘there’, *amb* ‘with’ or *i* ‘and’. Independently of the variable-invari-

able morphological dichotomy, words can be classified into lexical categories, also called *word classes*, depending on their syntactic and semantic properties.

2.1 Main word classes

Catalan grammar has traditionally distinguished between nine different word classes: nouns, adjectives, articles, verbs, adverbs, pronouns, prepositions, conjunctions and interjections (Fabra 1933; 1956). Modern linguistics has added the class of determiners, which includes articles – now considered a subclass – and various types of words such as demonstratives, unstressed possessives and quantifiers, which were previously included within the lexical category of adjectives. This change is justified by the fact that all these types of words share certain features, like prenominal position (*el pare* ‘the father’, *un pare* ‘a father’, *aquest pare* ‘this father’, *mon pare* ‘my father’, *algun pare* ‘some father’), in contrast with the usually postnominal position of qualifying adjectives (*pare bonadós* ‘good father’). We summarise the main semantic and syntactic features of these word classes below, in accordance with the most recent Catalan grammars (GCC; GIEC).

2.1.1 Nouns

Nouns denote countable or uncountable entities like persons (*home* ‘man’), animals (*gos* ‘dog’), objects (*cadira* ‘chair’), substances (*llet* ‘milk’), places (*ciutat* ‘city’) or ideas (*sorpresa* ‘surprise’). From the point of view of syntax, nouns can be the nucleus of a phrase – the noun phrase – which can act in a sentence as either the subject (*El nen menja* ‘The boy eats’), the predicate nominative after a linking verb (*El pare és el director* ‘The father is the director’), the object of a transitive verb (*Bevem llet* ‘We drink milk’) or the complement of a preposition (*Els llibres són al prestatge* ‘The books are on the shelf’). They can also act as the adjunct of another noun (*Barcelona, capital de Catalunya, és una gran ciutat* ‘Barcelona, capital of Catalonia, is a great city’).

2.1.2 Adjectives

From a semantic point of view, adjectives denote features which are attributed to the referent of a noun, as in *mare feliç* ‘happy mother’, *rosa roja* ‘red rose’ and *idea brillant* ‘brilliant idea’. Syntactically, adjectives can act as adjuncts to nouns, usually post-posed (*una nena eixerida* ‘a clever girl’, *una situació incòmoda* ‘an embarrassing situation’), and they can also serve as predicates after a copular verb (*L’àvia és sorda* ‘The grandmother is deaf’) or secondary predicates after a non-copular verb (*Hem arribat a casa cansats* ‘We arrived home tired’).

2.1.3 Determiners

Determiners initiate a noun phrase and indicate whether the noun phrase introduces either a new element in the speech, as in *M'ho ha dit un veí* 'A neighbour told [it to] me', or a known item, as in *M'ho ha dit el veí de dalt* 'The upstairs neighbour told [it to] me' or *M'ho ha dit aquest veí* 'This neighbour told [it to] me'. Within the determiner word class there are several subclasses, including articles, both definite (*el mestre* 'the teacher') and indefinite (*un mestre* 'a teacher'), demonstratives (*aquella casa llunyana* 'that distant house'), unstressed possessives (*ma mare* 'my mother') and certain quantifiers (*alguns estudiants* 'some students').

2.1.4 Verbs

From a semantic point of view, verbs express states, as in *Estimo el Daniel* 'I love Daniel', activities, as in *He llegit tota la tarda* 'I read all afternoon', achievements, as in *Han coronat el cim de l'Everest* 'They have reached the summit of Mount Everest', and accomplishments, as in *Ha après anglès amb molt d'esforç* 'He/She has learned English with great effort'. Syntactically, there are two main types of verb: predicative verbs, which constitute the nucleus of the predicate, as in *Menja massa* 'He/She eats too much', and copular verbs, which link the subject with its complement, as in *Aquest llibre sembla ben interessant* 'This book seems very interesting'.

2.1.5 Adverbs

Adverbs are linguistic elements that typically modify a verb, as we see in *Ens trobarem aquí* 'We will meet here', *Torna aviat* 'Come back soon' or *Parla clar* 'He/She speaks plainly'. Adverbs may also modify an adjective, as in *És ben baixa* 'She is very short', another adverb, as in *Viu molt lluny* 'He/She lives very far away', or a whole sentence, as in *Francament, no ho entenc* 'Frankly, I don't understand it'.

2.1.6 Pronouns

Pronouns are words whose meaning is defined on the basis of deictic or anaphoric relations. The deictic meaning of a pronoun depends on the particular speech act involved, so the referent of the pronoun can only be determined if we know who the speech act participants are and what the spatiotemporal context is. For example, in *Tu em dius això?* 'You_{SING} tell me that?', we cannot know what people are referred to by 'you' and 'me' or what it was that 'you' told 'me' in the absence of further information. Often the meaning of a pronoun is anaphoric, that is, it refers to information that

appeared earlier in the same utterance, as illustrated by the pronoun *hi* in the sentence *La Maria va anar a Grenoble i hi va trobar feina* ‘Maria went to Grenoble and found work there’. Syntactically, pronouns usually assume the same roles as nouns, as in *Vosaltres guanyeu* ‘You_{PL} win’, *Compra-ho* ‘Buy it’ or *La Maria va venir amb nosaltres* ‘Maria came with us’. However, they may also carry out an adjectival function, as in *Quina cançó t’estimes més?* ‘Which song do you prefer?’, or an adverbial function, as in *No hi vagis* ‘Don’t go there’. One subclass of pronouns are the so-called *personal pronouns*, which may refer to the participants in the speech act, either the addresser, as in *Ho dic jo* ‘It is I who say it’ and *Dona’ns-ho* ‘Give it to us’, or addressee, as in *Així tu vindràs?* ‘Then you_{SING} will come?’ and *Us esperen* ‘They are waiting for you_{PL}’, or they may refer to someone who is not directly involved in the speech act, as in *L’estimo* ‘I love him/her’ or *Aquest regal és per a ells* ‘This present is for them’.

2.1.7 Prepositions

Prepositions are words which are located at the head of a syntactic constituent, usually nominal, to form a prepositional phrase, and they express a kind of grammatical relationship between this constituent and another syntactic constituent in the sentence, as in *Ho farem en tres dies* ‘We will do it in three days’, *Ha comprat un conjunt de cadires* ‘He/She has bought a set of chairs’ and *Anem sense les maletes* ‘Let’s go without the suitcases’. Most prepositions are clitic, that is, they are unstressed and form a phonological unit with the word that follows them.

2.1.8 Conjunctions

Conjunctions connect either two words or two groups of words with the same syntactic function, as in *Menja pa i formatge* ‘He/She eats bread and cheese’ and *Vols uns pantalons estrets o unes faldilles amples?* ‘Do you want some tight trousers or some wide skirts?’. They can also join two coordinated or subordinate elements, as in *Prou que vindria, però no puc* ‘I’d love to come, but I can’t’ and *Diu que t’enyora* ‘He/She says that he/she misses you’. Like prepositions, some conjunctions are clitic and require the phonological support of another word.

2.1.9 Interjections

Finally, interjections are words that are syntactically equivalent to a sentence. They express pragmatic values associated with the speaker’s attitude, as in *Ep! Para compte, que em trepitges!* ‘Hey! Look out! You’re stepping on me’, *Uf! Quina feinada!* ‘Ugh! What hard work!’ or *Au, va! Qui s’ho creu, això?* ‘Sheesh! Who can believe that?’.

2.2 Transversal word classes

Sometimes words coming from different lexical categories have shared syntactic or semantic features that justify their being grouped together in what are known as *transversal word classes*. The seven transversal word classes are demonstratives, possessives, quantifiers, relatives, interrogatives, exclamatives and connectors. We will now summarise the main shared features that characterise each transversal word class, as set out in the most recent grammars of Catalan (*GCC*, *GIEC*).

First, demonstratives are words which have a deictic function. They introduce distinctions on the basis of the proximity or remoteness from the space occupied by the addresser and the addressee, as well as on the basis of elements that have appeared in the utterance. Demonstratives may be determiners, as in *aquest* *estoi*g ‘this case (near me)/that case (near you)’ versus *aquell* *estoi*g ‘that case (near him/her)’, pronouns, as in *vull* *això* ‘I want this (near me)/I want that (near you)’ versus *vull* *allò* ‘I want that (near him/her)’ or adverbs, as in *Vine aquí* ‘Come here’ versus *Ves allí* ‘Go there’ and *Fes-ho així* ‘Do it like this’.

The second transversal class are possessives, which are words that express a relationship of belonging. They are equivalent to the phrase *de* + personal pronoun, so that *el seu* *cotxe* ‘his/her car’ is equivalent to *el cotxe d’ell/d’ella* ‘the car of him/her’. Possessives can be determiners, as in *ta* *germana* ‘your sister’, as well as (prenominal or postnominal) adjectives, as in *el meu* *home* ‘my husband’, *un veí teu* ‘a neighbour of yours’, *casa nostra* ‘our house’, *és seu* ‘it is his/hers/theirs’ or *davant vostra* ‘in front of you_{PL}’.

With regard to quantifiers, these are words that express quantity, degree or number. They can be determiners, as in *Tenim quatre* *criatures precioses* ‘We have four beautiful children’ and *Cada* *participant va rebre un record* ‘Each participant received a memento’, pronouns, as in *Tothom* *l’estima* ‘Everybody loves him/her’ and *No em cal res* ‘I don’t need anything’, or adverbs, as in *Aquest vestit et queda més bé que l’altre* ‘This dress fits you better than the other one’ and *Ja hem caminat prou per avui!* ‘We have walked enough for today!’.

Moving on to relatives, these are words whose meaning is dependent on another element that has appeared previously in the utterance. They usually belong to the lexical category of pronouns, as in *el bolígraf que buscaves* ‘the pen that you_{SING} were looking for’, *el noi de qui es va enamorar* ‘the boy who she fell in love with’ or *el problema del qual et parlava* ‘the problem which I was telling you about’, but they can also be adverbs, as in *la universitat on treballa* ‘the university where he/she works’, or determiners, *Vol recomprar-ho, la qual cosa és impossible* ‘He/She wants to buy it back, which is impossible’.

As for interrogatives, these are words that initiate the sentence in which they are included and serve to formulate a question. They can be pronouns, as in *Què vols?* ‘What do you want?’, determiners, as in *Quins exercicis has fet?* ‘What exercises have you_{SING} done?’, or adverbs, as in *On anirem a parar?* ‘Where will we end up?’.

Exclamatives are also words that start the sentence in which they appear but in this case they add weight or emphasis to a particular notion. They can be determiners, as in *Quant de temps!* ‘What a long time!’, adverbs, as in *Com trona!* ‘How it is thundering!’, or pronouns, as in *Però què t’empatolles!* ‘But what are you_{SING} talking about!’.

The final transversal word class are connectors. This class contains both conjunctions – which, as we have seen, connect words, phrases or sentences – and so-called *parenthetical connectors*, whose function is to make connections within the text across sentences. Parenthetical connectors may be adverbs, as in *Conseqüentment*, *no podem aprovar la proposta* ‘Consequently, we cannot approve the proposal’, or locutions, as in *No obstant això*, *votarem a favor* ‘This notwithstanding, we will vote in favor’.

2.3 Closed and open classes

Some of the lexical categories that we have just seen are closed classes, made up of a finite and limited number of members. Conjunctions, prepositions, pronouns, determiners and adverbs – except adverbs of manner of the sort that end in *-ment* ‘-ly’ – are all closed classes, and their meaning is fundamentally grammatical in nature. So too are the article subclass and the transversal classes of demonstratives, possessives, non-numerical quantifiers, relatives, interrogatives, exclamatives and connectors. However, all the remaining lexical categories – nouns, adjectives, verbs and modal adverbs ending in *-ment* ‘-ly’ – are open classes, whose meaning is lexical rather than grammatical. The open classes contain a vast number of members and are productive, meaning that new members can readily be created and added to these classes. By the same token, words in open classes may fall out of use and be lost.

2.4 Variable and invariable classes

Some word classes such as adverbs, prepositions, conjunctions and interjections are invariable, or non-inflective, meaning that they cannot take inflections. By contrast, verbs, nouns, adjectives, determiners – including articles – and most of the pronouns are variable, or inflected. This can be seen in the verbal forms *canto* ‘I sing’, *cantàvem* ‘we sang’ and *cantaran* ‘they will sing’, in the noun forms *ull* ‘eye’ and *ulls* ‘eyes’, in the adjective forms *nen petit* ‘little boy’, *neneta petita* ‘little girl’, *nens petits* ‘little boys’ and *nenes petites* ‘little girls’, in the article forms *el llibre* ‘the book’, *la llibreta* ‘the notebook’, *els llibres* ‘the books’ and *les llibretes* ‘the notebooks’, and in the pronominal forms *ell* ‘he’, *ella* ‘she’, *ells* ‘they_{MASC}’ and *elles* ‘they_{FEM}’. The set of inflectional rules affecting the word class consisting of verbs is called *verbal inflection*, while *nominal inflection* refers to the set of inflectional rules that apply to nouns, adjectives, determiners and pronouns. Looking at just one of the many inflectional forms for verbs, the inflectional marker /d/ – orthographically rendered word-finally as *-t* as in

cantat ‘sung’, *perdut* ‘lost’ or *sentit* ‘heard’ – yields the past participle, which can play a syntactic role similar to an adjective and is itself affected by nominal inflection, as we see in *Vaig trobar un got escantellat* ‘I found a broken glass’ versus *Vaig trobar dues tasses escantellades* ‘I found two broken cups’. Both nominal and verbal inflection are further analyzed below, the sources for our descriptions being once again the most recent grammars of Catalan (GCC, GIEC).

3 Inflectional categories

As we have seen, inflection allows us to distinguish between invariable or non-inflective word classes and variable or inflective ones. Focusing on the latter, let us proceed to characterise the various inflections according to their grammatical properties, which are also named *inflectional categories*. The inflectional categories that we find in Catalan are gender, number, case, tense, aspect, mood, person and conjugation.

3.1 Nominal inflection

The main categories of nominal inflection in Catalan are gender and number. They are both binary: gender can be masculine or feminine, while number can be singular or plural. Besides gender and number, the pronominal system also shows some relics of the Latin case declension system. Finally, though person as a grammatical property affects personal pronouns and possessives, the expression of person in the nominal classes is – unlike what we see in verbs – not inflectional but rather lexical.

3.1.1 Gender

Gender is an inherent grammatical property of nouns. Thus, all nouns in Catalan belong to one of two large morphological classes, masculine nouns, like *sol* ‘sun’, *gos* ‘male dog’, *poema* ‘poem’, *roure* ‘oak’, *carro* ‘cart’, *bigoti* ‘moustache’ and *globus* ‘balloon’, and feminine nouns, like *roca* ‘rock’, *gossa* ‘female dog’, *mel* ‘honey’, *febre* ‘fever’, *moto* ‘motorcycle’ and *tribu* ‘tribe’. In most cases, gender has no semantic basis: there is no logical reason why *vi* ‘wine’ should be masculine while *llet* ‘milk’ is feminine. Nevertheless, in some cases gender is associated with semantic differences in sexed beings like humans or animals, as we see in *Aquest jove és mestre* ‘This young man is a teacher’ versus *Aquesta jove és mestra* ‘This young woman is a teacher’. Similarly, this gender distinction affects the third person pronoun, whether singular or plural: *ell* ‘he’ versus *ella* ‘she’, *ells* ‘they_{MASC}’ versus *elles* ‘they_{FEM}’, *mira’l* ‘look at him’ versus *mira-la* ‘look at her’ and *mira’ls* ‘look at them_{MASC}’ versus *mira-les* ‘look at

them_{FEM}'. As for animals, in only a few cases do gender distinctions reflect semantic differences linked to sex, as we see in *gat* 'male cat' versus *gata* 'female cat', *lop* 'male wolf' versus *loba* 'female wolf' or *os* 'male bear' versus *ossa* 'female bear'. By contrast, the grammatical gender of most animals, whether masculine like *poll* 'louse' and *hipopòtam* 'hippopotamus' or feminine like *formiga* 'ant' and *girafa* 'giraffe', has no semantic foundation.

Unlike nouns, all other nominal word classes – articles, other determiners, adjectives and participle verb forms – achieve gender through agreement with a noun. Thus, gender agreement (together with number agreement, as we will see) is a means of reflecting the syntactic relations between certain words within a phrase or sentence. For example, if we compare the phrase *aquell motiu secret* 'that secret motive' with *aquella raó secreta* 'that secret reason', we see agreement inside each noun phrase between the determiner and the noun and between the adjective and the noun. By the same token, in *un nen enfadat* 'an annoyed boy' versus *una nena enfadada* 'an annoyed girl', we see agreement between the indefinite article and the noun and between the participle and the noun. Agreement also occurs between a nominal predicate and a noun subject, as in *La Núria i la Mercè són molt llestes* 'Núria and Mercè are very clever'. Regarding past participles, these may establish agreement relations with other syntactic components, such as accusative clitic pronouns when the participle is part of a compound verbal tense, as we see in *On és la meva bossa? Que l'has vista?* 'Where is my bag? Have you seen it?'. Past participle agreement is also observable with the subject in passive constructions, as in *La carretera serà asfaltada properament* 'The road will be paved soon', and in absolute clauses, as in *Acabada la reparació, el lampista va marxar* 'The repairs being finished, the electrician left'.

3.1.2 Number

The grammatical property of number provides quantitative distinctions in nouns. Countable nouns – nouns referring to things which can be counted – have two inflected forms: singular, associated with the meaning 'one thing', and plural, associated with the meaning 'more than one thing'. The plural number inflection is -s, as seen in *una cullera* 'one spoon' versus *dues culleres* 'two spoons' or *cap got* 'no glass' versus *tots els gots* 'all the glasses'. Third person pronouns also inflect for number: *ell* 'he' versus *ells* 'they_{MASC}' and *ella* 'she' versus *elles* 'they_{FEM}', as well as *mira'l* 'look at him' versus *mira'ls* 'look at them_{MASC}' and *mira-la* 'look at her' versus *mira-les* 'look at them_{FEM}'.

Unlike nouns, other nominal word classes like articles and other determiners, adjectives and participles acquire number through agreement with a noun. Again, like gender agreement, number agreement is a way to show the syntactic relationships between certain words in a phrase or sentence. For example, comparing the noun phrases *un sofà còmode* 'a comfortable sofa' versus *uns sofàs còmodes* 'some comfor-

table sofas', we see agreement between the indefinite article and the noun and between the adjective and the noun, and likewise in *el coixí tacat* 'the stained pillow' versus *els coixins tacats* 'the stained pillows', agreement between the definite article and the noun and between the participle and the noun. Number agreement also occurs between nominal predicates and the noun subject, as in *El Jordi i el Roger són molt eixerits* 'Jordi and Roger are very clever'. Regarding past participles, these may show number and gender agreement with other syntactic components, such as accusative clitic pronouns, when the participle is part of a compound verbal tense, as we see in *No trobo les claus. Dec haver-les perdudes* 'I can't find my keys. I must have lost them'. Past participle agreement is also observable with the subject in passive constructions, as in *Els llibres han de ser retornats a la biblioteca dins del termini* 'Books must be returned to the library by the due date', and in absolute clauses, as in *Escoltats tots els arguments, la comissió va deliberar* 'All the arguments having been heard, the commission deliberated'.

Noun subjects also induce number agreement in verbs, together with person agreement, as we see in *Aquest estudiant treballa molt* 'This student works hard' versus *Aquests estudiants treballen molt* 'These students work hard' and *Vindrà una amiga meva* 'One of my friends will come' versus *Vindran unes amigues meves* 'Some of my friends will come'.

3.1.3 Case

Case is a grammatical property which serves to express differences related to the syntactic functions of nominal elements. In Catalan only a few personal pronouns show case distinctions, a vestige of their Latin origin. This is seen in the first person singular pronoun, which is *jo* 'I' when it acts as a subject, but *mi* 'me' when it acts as a prepositional complement. Note the contrast between *Li ho diré jo* 'I will tell him/her' and *Ja torneu a parlar de mi?* 'Are you talking about me again?' or *Vine amb mi!* 'Come with me!'. Thus, the first person singular pronoun has the nominative form *jo* but the oblique form *mi*, the difference between the two cases deriving from their separate Latin roots.

Similarly, the third person pronouns, in their weak or clitic form, show differences between their use as a direct object and their use as an indirect object, on this occasion with case differences expressed by means of inflection. The accusative forms are *el* 'him/it', *la* 'her/it', *els* 'them_{MASC}' and *les* 'them_{FEM}', with differences for gender and number, as we see in the following examples: *Busco el pare però no el trobo* 'I'm searching for my father but I can't find him', *Compraré una llibreta i la duré a classe* 'I'll buy a notebook and take it to class', *Si trobes els meus apunts, deixa'ls damunt de la taula* 'If you find my notes, leave them on the table' and *Quan acabis les postals, envia-les* 'When you finish the postcards, send them'. By contrast, the dative forms are *li* '(to) him/her' and *els* '(to) them', with differences in number but not gender: *Si veus*

el Daniel, dona-li la bona notícia ‘If you see Daniel, give him the good news’, *He parlat amb la Maria i li he explicat el problema* ‘I talked to Maria and told her the problem’, *Quan vagis a visitar els teus cosins, torna’ls aquest llibre, si us plau* ‘When you visit your cousins, give them this book back, please’ and *Quan vegis les teves germanes, dona’ls records de part meua* ‘When you see your sisters, give them my regards’. Note that on this occasion the case differences are expressed by means of inflection, since all these forms have the same root, /l/, which can receive – in addition to feminine and plural markers /a/ and /z/ respectively – the dative marker /i/.¹

3.1.4 Person

The grammatical property of person is tied to the speech act and specifically to the speech act participants. First person refers to the addresser, second person refers to the addressee and third person refers to any person that is not the addresser or the addressee, as can be seen in *Jo et parlo d’ell* ‘I(1st) am speaking to you(2nd) about him (3rd)’. Of the nominal elements, two subclasses are affected by person, personal pronouns and possessives. In these word classes, the person is not expressed by inflectional means, but instead it is the roots themselves that express the person. We can see this in the stressed personal pronouns *jo* ‘I’, *tu* ‘you_{SING}’, *ell* ‘he’, *ella* ‘she’, *nosaltres* ‘we’, *vosaltres* ‘you_{PL}’, *ells* ‘they_{MASC}’ and *elles* ‘they_{FEM}’, and also in the unstressed personal pronouns *em* ‘me’, *et* ‘you_{SING}’, *el* ‘him’, *la* ‘her’, *ens* ‘us’, *us* ‘you_{PL}’, *els* ‘them_{MASC}’ and *les* ‘them_{FEM}’. Person inflections are also apparent in possessives, whether stressed, like *meu* ‘my/mine’, *teu* ‘your/yours_{SING}’, *seu* ‘his/her/hers/their/theirs’, *nostre* ‘our/ours’, *vostre* ‘your/yours_{PL}’ and *llur* ‘their’, or unstressed, though in the latter case only the singular and third person plural forms exist, like *mon* ‘my’, *ton* ‘your_{SING}’ and *son* ‘his/her/their’.

3.2 Verbal inflection

The categories of verbal inflection in Catalan are as follows: conjugation, with three classes; tense, with a triple distinction between present, past and future; aspect, a binary category which distinguishes perfective tenses from imperfective ones; mood, which can be indicative, subjunctive or imperative; person, which involves a distinction between first, second and third person; and number, another binary category, which distinguishes between singular and plural.

¹ In Standard Catalan, the dative marker is /i/ in singular and Ø in plural, as we see in *li dic adeu* ‘I tell him/her goodbye’ versus *els dic adeu* ‘I tell them goodbye’, though in some colloquial varieties of Catalan the plural may also receive the /i/ marker: *els hi dic adeu* ‘I tell them goodbye’.

3.2.1 Conjugation

Conjugation is an inherent grammatical property of verbs by which they can be separated on a purely formal basis into various classes, or *conjugations*, each of them obeying a particular inflectional pattern. In Catalan there are three such classes: first conjugation verbs, with the infinitive ending in *-ar* (stressed), like *cantar* ‘to sing’; second conjugation verbs, with the infinitive predominantly ending in *-re* (unstressed), like *perdre* ‘to lose’, but also in *-er* (unstressed), like *témer* ‘to fear’, in *-er* (stressed), like *poder* ‘can’, and in *-r*, like *dur* ‘to carry’; and third conjugation verbs, with the infinitive ending in *-ir* (stressed), like *dormir* ‘to sleep’.

Within the third conjugation there are two subclasses, the so-called *third pure class* and *third inchoative class*, this latter name having a historical justification rather than a grammatical one. In Latin, inchoative verbs express the beginning of an action or a change of state through the infix *-ESC-*, as in *CĂLEŌ* ‘to be warm’ versus *CĂLESCO* ‘to become warm’. In Catalan, the aspectual value of this infix has been lost, but there exist certain verbs that, in some inflected forms, receive a root extension derived from *-ESC-*, though today it is void of meaning (Pérez Saldanya 1998). Compare, for instance, the third person singular present indicative *dorm* ‘he/she sleeps’, from *dormir* ‘to sleep’, a third pure conjugation verb, and *parteix* ‘he/she divides’, from *partir* ‘to divide’, a third inchoative conjugation verb.

3.2.2 Tense

The grammatical property of tense is deictic, because it locates the action expressed by the verb in relation to the moment of the speech act. Tense may be present, past or future. The present tense expresses simultaneity relative to the speech act, as in *No cridis, que dorm* ‘Don’t shout, he/she is sleeping’, while the past tense expresses anteriority, as in *Ahir va ploure tot el dia* ‘Yesterday it rained all day’, and the future tense expresses posteriority, as in *Vindrem la setmana entrant* ‘We will come next week’.

3.2.3 Aspect

The grammatical property of aspect is related to the internal temporality of situations denoted by the predicate. In Catalan, only one aspectual distinction has an inflectional expression, and it is the binary opposition between perfective and imperfective aspect. The perfective aspect presents the situation denoted by the predicate as a situation which is closed or finished, as in *L’any passat va acabar la carrera* ‘Last year he/she finished his/her university degree’ and *Ahir em vaig trobar amb la Teresa i vam fer-la petar una bona estona* ‘Yesterday I met Teresa and we chatted a long time’. By

contrast, the imperfective aspect presents the situation as being in progress, as in *Quan el corrector va entrar al despatx, el director parlava per telèfon* ‘When the proof-reader came into the office, the editor was talking on the phone’, or habitual, as in *Jo, mentre estudiava a la universitat, treballava en una caixa* ‘While I was studying at university, I worked in a bank’.

3.2.4 Mood

Verbal mood serves to introduce distinctions linked to the attitude of the speaker with respect to what he/she is saying. In Catalan there exist three moods, indicative, subjunctive and imperative. The indicative mood appears in assertive contexts, as in *Plou* ‘It is raining’ and *El Daniel explica un conte al nens* ‘Daniel is telling a story to the children’. By contrast, the subjunctive mood is linked to non-assertive contexts. Thus, the subjunctive is the mood of possibility, desire or counterfactuality, as we see in sentences like *Podria ser que plogués* ‘It might be raining’, *Tant de bo que plogués!* ‘I wish it would rain!’ and *Si no plogués, podriem fer un tomb* ‘If it weren’t raining, we could go for a walk’. Finally, the imperative mood is used to express requests or orders directed at the addressee, as in *Deixa-ho córrer* ‘Let it go’ and *Tanca la boca!* ‘Shut your mouth!’.

3.2.5 Person

We have referred to the grammatical property of person above (3.1.4). It is expressed in verbal forms – together with number – through agreement with the subject. The verb tenses thus have six inflected forms to satisfy this agreement, as exemplified by the present indicative tense of *cantar* ‘to sing’: first person: *canto* ‘I sing’, *cantem* ‘we sing’; second person: *cantes* ‘you_{SING} sing’, *canteu* ‘you_{PL} sing’; and third person: *canta* ‘he/she sings’, *canten* ‘they sing’.

3.2.6 Number

We have already discussed the grammatical property of number above (3.1.2). With regard to number inflections as applied to verbs, it is worth noting that the singular/plural dichotomy applies within each grammatical person, as can be seen in the inflections of *cantar* ‘to sing’ shown in 3.2.5.

4 Paradigms

All the inflective forms of a given word constitute its inflectional *paradigm*. Thus, the paradigm of the noun *dona* is *dona*, *dones* ‘woman, women’ and the paradigm of the adjective *dur* ‘hard’ is *dur*, *dura*, *durs*, *dures*. The full paradigm of a verb is much more complex than a nominal paradigm. For example, *canto*, *cantaves*, *cantarà*, *cantéssim*, *vau cantar* and *han cantat* are only a few of the many conjugated forms of *cantar* ‘to sing’, comprising – like most verbs – some 50 simple forms and 68 compound forms. Given the complexity of the verbal conjugation, the term *paradigm* may also refer to a small part of the full verbal paradigm, usually the six forms of a given tense. For instance, the future tense paradigm of *cantar* ‘to sing’ is *cantaré* ‘I will sing’, *cantaràs* ‘you_{SING} will sing’, *cantarà* ‘he/she will sing’, *cantarem* ‘we will sing’, *cantareu* ‘you_{PL} will sing’ and *cantaran* ‘they will sing’.

Each one of the forms of an inflectional paradigm corresponds to the same lexical word. Therefore, for practical purposes it is necessary to agree upon just one of those forms to serve as the reference form under which, for example, the word can be found in a dictionary. For nouns, the reference form is the singular. Thus, for instance, *taula* ‘table’ represents both the singular *taula* ‘table’ and the plural *taules* ‘tables’. For all other nominal classes, the reference form is the masculine singular, so that the form *gros* ‘big’ represents the masculine singular *gros*, the feminine singular *grossa*, the masculine plural *grossos* and the feminine plural *grosses*. Finally, the reference form for verbs is the infinitive. For example, *estimar* ‘to love’ represents all the numerous inflected forms of this verb.

4.1 Markedness oppositions

Inflection is often expressed through inflectional markers added after the stem, but it can also be realised through the absence of markers, as we see in the masculine singular form of the adjective *fort* ‘strong’ (which is the reference form, as noted above). By contrast, the feminine plural form of the same adjective, *fortes*, has both a feminine marker (-e-) and a plural marker (-s). The distribution within paradigms between forms with and without inflectional markers is directly related to the oppositions that exist within inflectional categories, and to the fact that inside these oppositions (which are usually binary, like masculine/feminine) one of the terms of the opposition is functionally marked (and often marked formally by an inflectional affix) and the other is functionally unmarked (and typically also unmarked formally, thus lacking an inflectional affix). Take, for example, the masculine/feminine opposition within the category of gender: the singular masculine form *fort* lacks inflectional affixes and is regarded as the unmarked form, while the singular feminine *forta* has the affix -a and is regarded as the marked form.

Unmarked forms generally show a larger distribution than marked forms and can be used in a generic sense that subsumes the marked form. Thus, with regard to gender, the masculine unmarked term in broad contexts includes the feminine one. Similarly, in number opposition, the singular is the unmarked term and can implicitly include the plural. For instance, in the sentence *El gat és un bon animal de companyia* ‘The cat makes a good pet’, the singular masculine inflected form *gat* ‘cat’ is used generically to indicate all members of this category, that is, all cats, whether male or female, yet it displays no inflectional markers, either for gender or for number.

Although less common than the binary dichotomy, an inflectional category can express itself through three terms, as we see in the case of person (first, second or third) and tense (past, present or future). Nonetheless, even in these cases, one of the three terms is unmarked relative to the others, and the relationships between the three terms can be reduced to binary expressions. Within the grammatical property of person, for instance, persons are either the subjective person (first person) or the non-subjective person (second or third person), and they are either being addressed (second person) or not (first person or third person). The unmarked person is therefore the third person, sometimes called the *non-person*, and proof of this is the fact that it can replace the first and/or second person in certain contexts. For example, a mother addressing her son can say *La mama t’estima un munt* ‘Mum loves you a lot’ or *Què li diu aquest nen guapo a la seva mama?* ‘What does this handsome boy say to his mother?’.

Similarly, within the grammatical property of tense, the past tense can be binarily opposed to all non-past tenses and the future tense to all non-future tenses. On the other hand, the present is the tense of the speech act itself and can therefore be used with a kind of universal “default” value. For instance, it can be employed in a past context, as in *Ahí me’l trobo assegut a la meva cadira i jo que li dic: Què? Estàs còmode, noi?* ‘Yesterday I find him sitting in my chair so I tell him: Well, are you comfortable, buddy?’, or in a future context, as in *Ves a dormir d’hora, que demà tens molta feina* ‘Go to bed early, since you have a lot of work tomorrow’. It is even used to express a kind of timeless meaning, as in *La Terra gira al voltant del Sol* ‘The Earth revolves around the Sun’.

These markedness oppositions can be illustrated by comparing the verbal form *perd* ‘he/she loses’, third person singular present indicative tense, which has neither person nor tense markers, with *perds* ‘you_{SING} lose’, second person singular present indicative tense, which has a person marker – the final *s* – but no tense marker, or *perdràs* ‘you_{SING} will lose’, second person singular future tense, which has both a person marker – the final *s* – and a tense marker – the cluster *rà*.

4.2 Defectiveness

Defectiveness consists of the absence within a paradigm of one or more of the usual inflective forms. For instance, the verb *ploure* ‘to rain’ has only third person forms, as in *Avui plou molt* ‘Today it is raining a lot’, logically enough given its impersonal character. Nouns may also be regarded as defective if they have plural forms but no singular forms, even though in such cases they usually refer to a single object, as in *unes tenalles* ‘some (= a pair of) pliers’.

4.3 Syncretism

Syncretism is the formal coincidence between two inflected forms that differ in grammatical specification. For instance, the noun *llapis* ‘pencil’ has the same form whether it is singular or plural, as we see in *un llapis* ‘one pencil’ versus *dos llapis* ‘two pencils’. Hence the noun *llapis* ‘pencil’ is said to be syncretic (or invariable) with respect to number.² The adjective *feliç* ‘happy’ is syncretic in the singular, because the masculine form, as in *un noi feliç* ‘a happy boy’, is the same as the feminine, *una noia feliç* ‘a happy girl’. The plural, on the other hand, is not syncretic: *uns nois felïços* ‘(some) happy boys’ versus *unes noies felïces* ‘(some) happy girls’. Different varieties of Catalan may vary in this regard. For example, some nominal elements are not syncretic in the singular in Western Catalan, where unstressed *a* and *e* are pronounced differently, but they are syncretic in Eastern Catalan, where unstressed *a* and *e* have been conflated into a schwa (⌈8.1 Dialects). Thus, *mestre* ‘male teacher’ and *mestra* ‘female teacher’ and *negre* ‘black_{MASC}’ and *negra* ‘black_{FEM}’ are pronounced differently in Western Catalan ([ˈmestre]/[ˈmestra], [ˈnɛɣre]/[ˈnɛɣra]) but pronounced identically in Eastern Catalan ([ˈmɛstrə], [ˈnɛɣrə] or [ˈnəɣrə]).³

In Catalan verbal paradigms, one typically finds syncretism between the first and third person singular forms in the imperfect indicative, as in *cantava* ‘(I/he/she) sang’, the conditional, as in *cantaria* ‘(I/he/she) would sing’, the present subjunctive, as in *(volen que) canti* ‘(they want) me/him/her to sing’ and the imperfect subjunctive, as in *(volien que) cantés* – or *(volien que) cantara* in Valencian – ‘(they wanted) me/him/her to sing’. Syncretism often occurs elsewhere in verbal paradigms, such as in the first and second person plural forms of the present indicative and present subjunctive: *dormim* ‘(we) sleep’/(*vol que*) *dormim* ‘(he/she wants) us to sleep’, and *dormiu* ‘(you_{PL}) sleep’/(*vol que*) *dormiu* ‘(he/she wants) you_{PL} to sleep’.

² That said, some colloquial varieties of Catalan use the analogical plural form *llapissos* ‘pencils’.

³ Regarding variation in the pronunciation of stressed vowels among Catalan dialects, ⌈8.1 Dialects.

5 Morphs, allomorphs and zero morph

Thus far we have dealt with morphemes from the point of view of their meaning and grammatical content. We must not forget, however, that morphemes are also linguistic elements with phonological content. Thus, the word *fort* ‘strong’ is based on the existence of a root with which several types of linguistic content are associated, namely the phonological form /'fɔrt/, the meaning ‘physically or mentally vigorous’ and some grammatical information such as the fact that it belongs in the word class of adjectives.

The phonological form of a morpheme, whether root or affix, is called a *morph*. It should be noted, however, that the relation between morphemes and morphs is not a simple one-to-one relationship. For this reason, we need to introduce the concepts of *zero morph*, *allomorphy* and *fusion*.

5.1 Zero morph

As we have seen, inflection is often expressed in Catalan through inflectional markers added after the stem. However, this can also be achieved through the absence of markers. Thus, the adjectival root /'fɔrt/ ‘strong’ receives neither gender marker nor number marker to become the word *fort* in its masculine singular inflected form, while the feminine forms receive the gender marker /a/, orthographically represented by *a* in the singular *forta* and *e* in the plural *fortes*, and plural forms receive the number marker /z/, orthographically *s*, hence *forts*, *fortes*. When a situation like this occurs, a morphological analysis can be proposed that postulates the existence of a morph consisting of no phonological form, called *zero morph* and represented by $\emptyset$ (Mascaró 1986). Therefore, in the previous example, the masculine singular form *fort* would be morphologically segmented as /'fɔrt+ $\emptyset$ + $\emptyset$ /, the feminine singular form *forta* as /'fɔrt+a+ $\emptyset$ /, the masculine plural form *forts* as /'fɔrt+ $\emptyset$ +z/ and the feminine plural form *fortes* – the only one without a zero morph – as /'fɔrt+a+z/.

5.2 Allomorphy

Sometimes one morpheme can have more than one phonological form, that is, it can be related to more than one morph depending on the grammatical context. Each of the morphs related to the same morpheme in different grammatical contexts is called an *allomorph* and the phenomenon by which this occurs is called *allomorphy*. For instance, the Catalan future marker can be /'ra/, /'re/ or /'rɛ/ depending on the grammatical person and number where this marker appears as well as the variety of Catalan being spoken. In all Catalan dialects the future marker is expressed by the allomorph /'ra/ in the second and third person singular and the third person plural, as

in *cantaràs* ‘you_{SING} will sing’, *cantarà* ‘he/she will sing’ and *cantaran* ‘they will sing’. The first person singular has the marker /'re/ in most of the Catalan linguistic domains, as in *cantaré* ‘I will sing’, although in some conservative northwestern varieties, especially in the Pallarese region, the marker is /'rɛ/ ([kanta'rɛ]), reflecting an earlier stage in the evolution of Romance [aj], whereby the stressed vowel has been closed one degree by the palatal glide yod before being lost: CANTAR(E) HA(B)EO > [kanta'rajo] > [kanta'raj] > *canta*['rɛ] (> *canta*['re]), MA(G)IS > ['majs] > ['mɛs] (> ['mes]) *més* ‘more’ and LACTE > ['lajte] > ['ʎɛ(j)t] (> ['ʎɛ(j)t]) *llet* ‘milk’ (Veny ¹²1998). Regarding the first and second person plural, *cantarem* ‘we will sing’ and *cantareu* ‘you_{PL} will sing’, the plural marker is /'re/ in Western Catalan, while Eastern Catalan exhibits the pronunciations /'rə/ in most of the Balearic varieties, /'rɛ/ in Central Catalan and some Balearic varieties, /'rɛ/ in Northern Catalan and, once again, /'re/ in Alguerese, due to the historical origin of this vowel, a Latin long E (HABĒMUS, HABĒTIS), the path of whose evolution in Catalan dialects was first described by Fabra in the early 20th century (1981 Dialects).

Moreover, in a few cases inflection involves the allomorphic variation of roots. One example of this is the irregular verb *voler* ‘to want’. The verbal forms *vull* ‘I want’ and *vol* ‘he/she wants’, the first and third person singular present indicative respectively, have no inflectional markers at all, but they are easily distinguished thanks to the allomorphic variation of the root, which selects the allomorph /vuʎ/ in the first person singular present indicative and the allomorph /vɔʎ/ in all the other present indicative persons.

5.3 Fusion

In the Catalan verbal inflection system we frequently encounter the phenomenon called *fusion*, whereby two or more grammatical properties are expressed through a single phonological form. Thus, in Catalan the grammatical properties of person and number always appear merged; the morph /z/ expresses at one and the same time the second person and the singular, as in *dormies* ‘you_{SING} slept’, while the morph /w/ indicates the second person together with the plural, as in *dormieu* ‘you_{PL} slept’. In Catalan we also find fusion between the expression of tense, aspect and mood. In the inflected form *cantàvem* ‘we sang’ the morph /va/ provides three kinds of grammatical information, namely that the verb is in the past tense, the imperfective aspect and the indicative mood, and similarly in the form (*volien que*) *cantéssim* ‘(they wanted) us to sing’ the segment /si/ informs us that the verb is in the past tense, the imperfective aspect and the subjunctive mood. For this reason, Catalan grammars usually name *person markers* inflectional affixes which express both person and number, just as inflectional affixes which simultaneously indicate tense, aspect and mood are referred to as *tense markers*. In addition, it is common to refer to six verbal persons rather than three: persons 1, 2 and 3 are the first, second and third person singular, while persons

4, 5 and 6 are the first, second and third person plural. (We will resort to this system on occasion here.)

6 Inflectional markers in nominal word classes

Catalan nominal words can be broken down into three components that always follow the same order, as illustrated in Figure 1: stem, gender marker and number marker.

Inflected word	stem	gender marker	number marker
<i>petit</i> , masculine singular	/pe'tit/	∅	∅
<i>petita</i> , feminine singular	/pe'tit/	/a/	∅
<i>petits</i> , masculine plural	/pe'tit/	∅	/z/
<i>petites</i> , feminine plural	/pe'tit/	/a/	/z/

Figure 1: Morphological segmentation of the inflected forms of the adjective *petit* ‘little’.

In this section we examine in detail the Catalan inflectional markers that affect the nominal word classes, that is, gender and number markers. We will exemplify nominal inflection mostly with reference to the inflected forms of nouns and adjectives, but we will also refer to some of the more common nominal closed classes such as articles, demonstratives, possessives and personal pronouns. We base our description on Standard Catalan; this variety is in fact very respectful of and therefore inclusive of the dialectal diversity we find in the Catalan domain (GIEC; Badia i Margarit 1994).

6.1 Inflectional gender markers

With regard to the expression of gender in Catalan, the main pattern is the absence of a gender marker to indicate masculine gender while the feminine gender is expressed by the marker *-a* in the singular and *-e-* in the plural (73 Spelling). Thus, for example, the roots /'un/ (indefinite article), /'noj/ ‘young person’ and /'alt/ ‘tall’ require no gender markers to become masculine inflected words in *un noi alt* ‘a tall boy’ and *uns nois alts* ‘(some) tall boys’, but do require a gender marker to become feminine inflected words in the phrases *una noia alta* ‘a tall girl’ and *unes noies altes* ‘(some) tall girls’. Phonologically and historically there exists a single feminine marker, /a/, although it exhibits two manifestations due to a regular phonetic law (Badia i Margarit 1981; Moll 1991) according to which A was maintained in absolute final position in Preliterary Catalan, e.g. ROSA > *rosa* ‘rose’, but -AS became -es due to the closed character of this syllable, i.e. ROSAS > *roses* ‘roses’ (710 Early Medieval Catalan for an

outline of historical grammar). Western Catalan has maintained this distinction and the feminine marker is pronounced differently in singular, *rosa* [ˈrɔza] (with the dialectal variant [ˈrɔzɛ]), and plural, *roses* [ˈrɔzɛs] (§8.1 Dialects). Catalan orthography, standardized at the beginning of the 20th century, is respectful of this Western pronunciation and the Old Catalan spelling tradition, although in Eastern Catalan the unstressed vowels *a* and *e* have been conflated into schwa ([ə]) and for this reason the feminine marker is pronounced identically in singular and plural, *rosa* [ˈrɔzə] and *roses* [ˈrɔzəs] (§14.1 The Language Reform, the Institut d'Estudis Catalans and the Work of Pompeu Fabra).

From what we have seen thus far, it follows that feminine nouns and feminine inflected forms of adjectives and other nominal elements usually end in *-a* in singular and *-es* in plural, while masculine nouns and masculine inflected forms of adjectives and other nominal elements typically end either in a consonant, as in *fred* 'cold', *cec* 'blind', *nas* 'nose', *all* 'garlic', *bony* 'lump', *ferrer* 'blacksmith', *malalt* 'ill', *aquest* 'this' and *hort* 'vegetable garden', or in a glide, as in *dau* 'dice', *meu* 'my/mine', *pou* 'well', *lacai* 'lackey', *remei* 'remedy' and *cofoi* 'satisfied'.⁴ Some masculine forms appear to end in a vowel, such as *pa* 'bread', *redó* 'round' and *ase* 'donkey', but in phonological terms their roots actually end in *-n*, /ˈpan/, /reˈdon/, /ˈazen/, as is evident in their feminine and/or plural forms, *pans* 'breads', *redona*, *redons*, *redones* 'round' and *àsens* 'donkeys' (which has become *ases* in Eastern Catalan by analogy, see 6.2), as well as in derived words, such as *panarra* 'someone who eats a lot of bread', *arredonir* 'to make round' and *asenet* 'little donkey'. In other cases, masculine forms display a final *-e* which must be explained as a support vowel added to avoid consonant clusters not allowed in final position. For instance, the root /ˈagr/ 'sour' takes the support vowel *-e* to become the inflected masculine forms *agre* (/ˈagr+Ø+Ø/) and *agres* (/ˈagr+Ø+z/), but this vowel reappears neither in the inflected feminine forms *agra* (/ˈagr+a+Ø/) and *agres* (/ˈagr+a+z/) nor in derived forms like *agror* 'sourness' (/ˈagr+'or/), even if it is also required in compounds like *agredolç* 'bittersweet'_{MASC-SING} (/ˈagr+'dols/) and *agredolça* 'bittersweet'_{FEM-SING} (/ˈagr+'dols+a/).

Beyond these main patterns, there exists a minority class of masculine nouns with the ending *-a*, as in *dia* 'day', *mapa* 'map' and *poeta* 'poet', which turns into *-e* in plural, as in *dies* 'days', *mapes* 'maps' and *poetes* 'poets'. By the same token, there is a minority class of feminine nouns without a gender marker, as in *por* 'fear', *sal* 'salt' and *serp* 'snake', which, when necessary, receive the support vowel *-e*, as in *febre* 'fever' and *torre* 'tower'. This support vowel can be converted into *a* by

⁴ In spoken Catalan some of these consonants can become mute. Thus, for example, the final *-r*, as in *fuster* 'cabinetmaker', is only pronounced in Valencian varieties, and the final occlusive of the consonant clusters *-lt*, *-nt* and *-mp*, as in *alt* 'tall', *pont* 'bridge' and *camp* 'field', is preserved in consecutive dialects but not in constitutive ones (§8.1 Dialects).

analogy in Western Catalan pronunciation, as in *llebre* ['ləβra] 'hare'. In rare cases, nouns can have other endings, as in the masculine forms *amo* 'owner', *frare* 'friar' – where the *e* cannot be explained as a support vowel – and *hebreu* 'Hebrew', and in the feminine forms *mare* 'mother' – the *e* again not justifiable as a support vowel – *ràdio* 'radio' and *tribu* 'tribe'. It is worth mentioning that masculine nouns ending in *-o* often come from Old Catalan nouns which initially had a support vowel *-e* which became *-o* for morphological reasons, a change which has only affected masculine nouns like *ferro* 'iron' (Old Catalan *ferre* < FĒRRU) and *monjo* 'monk' (Old Catalan *monge* < Vulgar Latin MONĬCU) (Batlle et al. 2016).

Thus, the gender opposition between masculine and feminine forms in nouns, adjectives and other nominal elements is usually manifested through the alternation between the zero morph in masculine forms and the marker /a/ in feminine ones. Nevertheless, gender opposition is sometimes expressed in the alternation between the endings *-e* in masculine and *-a* in feminine. This alternation is found when the nominal root ends in a consonant cluster that needs the support vowel *-e* in masculine forms, as we see in /'sɔgr/ 'parent-in-law' and /'pɔbr/ 'poor', and is externally visible in singular forms, as in *sogre* 'father-in-law' versus *sogra* 'mother-in-law' and *pobre* 'poor_{MASC}' versus *pobra* 'poor_{FEM}'. This spelling difference reflects a phonetic distinction in Western Catalan, ['sɔɣre] versus ['sɔɣra] and ['pɔβre] versus ['pɔβra], but not in Eastern Catalan, due to the conflation of unstressed *e* and *a* in schwa that we noted above. By contrast, in plural forms the nature of the vowel is not externally distinguishable, because the gender marker *-a* becomes *-e-* due to the closed character of the syllable, and then it coincides with the support vowel present in masculine forms. We see this in *sogres* 'parents-in-law' and *pobres* 'poor_{PL}', which correspond to the patterns /'sɔgr+Ø+z/ and /'pɔbr+Ø+z/ in the masculine plural forms and /'sɔgr+a+z/ and /'pɔbr+a+z/ in the feminine plural ones, pronounced ['sɔɣres] and ['pɔβres] in Western Catalan and ['sɔɣrəs] and ['pɔβrəs] in Eastern Catalan. Finally, in sporadic cases the gender opposition is manifested through the alternation between the endings *-o* in masculine and *-a* in feminine, as in *gitano* 'male gypsy' versus *gitana* 'female gypsy', and *-u* ([w]) in masculine and *-a* in feminine, as in *ateu* 'male atheist' versus *atea* 'female atheist'. Note that in all cases, when gender opposition is marked by inflectional means, the feminine form always ends in *-a* while the masculine one usually exhibits the zero morph or, less often, the endings *-e*, *-o* or *-u* ([w]).

Although most Catalan adjectives are gender-variable and exhibit the alternations described above, there also exist invariable adjectives. The most frequent endings for these gender-invariable adjectives are as follows, all of them with a final consonant: *-al*, as in *lleial* 'loyal', *-ant*, as in *brillant* 'shiny', *-ar*, as in *vulgar* 'vulgar', *-el*, as in *rebel* 'rebellious', *-ent*, as in *diferent* 'different', *-il*, as in *humil* 'humble' and *-(i)or*, as in *major* 'main'. One notable case of gender invariance is evident in adjectives ending in *-aç*, *-oç* and *-iç*, such as *eficaç* 'effective', *veloç* 'fast' and *feliç* 'happy'. These adjectives have a single singular inflected form which serves for both

masculine and feminine, as we see in *un remei eficaç* ‘an effective remedy’ and *una solució eficac* ‘an effective solution’. However, these words have two inflected plural forms, one for the feminine and another for the masculine, as seen in *uns remeis eficaços* ‘(some) effective remedies’ versus *unes solucions eficaces* ‘(some) effective solutions’. Historically these vowels are support vowels inserted to make possible the syllabification of the plural marker /z/ after a root ending in /s/. The vowel *e* present in the feminine plural form is the usual support vowel in Catalan, but it coincides with the realisation -e- in plurals of the feminine marker /a/. The vowel *o* present in the masculine plural form is a support vowel which is morphologically conditioned, because it only appears in masculine forms, and is the transformation of a historical *e*. For instance, the noun root /'bras/ ‘arm’ had the inflected forms *braç* ‘arm’ and *braces* ‘arms’ in Old Catalan, the latter with the usual support vowel *e*, but in Modern Catalan the plural has become *braços* ‘arms’, presumably due to the masculine gender of this noun.

In addition to the endings seen above, in Catalan there exist several gender-invariable adjectives with the final vowel -e or, less often, -a, such as *alegre* ‘cheerful’, *enorme* ‘enormous’, *jove* ‘young’, *lliure* ‘free’ and *nòmada* ‘nomadic’, pronounced in both cases [ə] in Eastern Catalan. Other common endings for gender-invariable adjectives are -ble, as in *feble* ‘weak’, -aire, as in *rondinaire* ‘grumbler’ and -ista, as in *anarquista* ‘anarchist’. In Western Catalan the gender-invariable adjectives ending in -e or in -a tend to shift by analogy into the variable adjective class with the alternation of -e in the masculine form and -a in the feminine: *un noi amable* ‘a kind boy’ versus *una noia amabl[a]* ‘a kind girl’, *un home xerraire* ‘a talkative man’ versus *una dona xerrair[a]* ‘a talkative woman’, *un partit anarquist[e]* ‘an anarchist party’ versus *una organització anarquista* ‘an anarchist organization’.

6.2 Inflectional number markers

The inflected nouns *casa* ‘house’ and *cases* ‘houses’ have a number morpheme, singular in *casa*, with the meaning ‘one element’, and plural in *cases*, with the meaning ‘more than one element’. In spite of this, only the plural has an inflective marker with phonological content, /z/, the final -s, while the singular is expressed through the absence of an inflective marker, that is, by a zero morph. This rule holds for the vast majority of nouns, as in *un rellotge antic* ‘an old clock’ versus *uns rellotges antics* ‘(some) old clocks’ and *la camisa nova* ‘the new shirt’ versus *les camises noves* ‘the new shirts’.

It should be noted that in Catalan there exists a group of nouns and adjectives whose roots end in an *n* that becomes mute in absolute final position. Thus, *mà* ‘hand’ (< MANU) has the root /'man/ and *comú* ‘common’ (< COMMŪNE) has the root /ko'mun/, just as *home* ‘man’ (< HŌMĪNE) has the root /'ɔmen/ and *orfe* ‘orphan’ (< ŌRPĤĀNU) has the root /'ɔrfen/. When these nouns and adjectives are oxytone or

monosyllabic, the nasal is kept in the inflected plural form.⁵ Thus the plural of *mà* ‘hand’ is *mans* ‘hands’ and the plural of *comú* ‘common_{MASC-SING}’ is *comuns* ‘common_{MASC-PL}’. When these nouns and adjectives are paroxytone, in Western Catalan – like in Old Catalan – the nasal is kept in the inflected plural form, but in Eastern Catalan this consonant has disappeared by analogy with the singular inflected form. Thus, the plural of *home* ‘man’ is *hòmens* or *homes* ‘men’ and the plural of *orfe* ‘orphan_{MASC-SING}’ is *òrfens* or *orfes* ‘orphan_{MASC-PL}’. Note that in all cases the feminine inflected forms contain this *n*, as in *comuna* ‘common_{FEM-SING}’, *comunes* ‘common_{FEM-PL}’ and *òrfena* ‘orphan_{FEM-SING}’, *òrfenes* ‘orphan_{FEM-PL}’, as do derived words, as in *maneta* ‘handle’, *manetes* ‘handyman’ and *manat* ‘handful’ or *homenot* ‘man (pejorative)’, *homenia* ‘virility’ and *homenatge* ‘homage’.

We also need to emphasise that in Catalan there exists a group of masculine nouns whose roots end in a sibilant, represented orthographically as *-s* ([s]), *-ç* ([s]), *-x* ([ks]), *-ix* ([ij]/[j]), *-tx* ([tʃ]) or *-ig* ([tʃ]), or in a sibilant followed by an occlusive, orthographically *-sc* ([sk]), *-st* ([st]), *-xt* ([kst]) or *-sp* ([sp]). This group constitutes a special case regarding number inflection, because in Catalan the plural marker is also a sibilant, /z/. Usually in these masculine nouns the plural marker is preceded by the support vowel *o*, which is morphologically conditioned because it only appears in masculine forms.⁶ Here are some examples with each one of these root endings: *pis* ‘flat’, *pisos* ‘flats’; *nas* ‘nose’, *nassos* ‘noses’; *estruç* ‘ostrich’, *estruços* ‘ostriches’; *annex* ‘attachment’, *annexos* ‘attachments’; *greix* ‘fat’, *greixos* ‘fats’; *cartutx* ‘cartridge’, *cartutxos* ‘cartridges’; *assaig* ‘essay’, *assajos* ‘essays’; *basc* ‘Basque’, *bascos* ‘Basques’; *arbus* ‘bush’, *arbustos* ‘bushes’; *text* ‘text’, *textos* ‘texts’; and *gesp* ‘lawn grass’, *gespos* ‘lawn grasses’. Similarly, the inflected masculine plural forms of adjectives whose roots end in a sibilant or sibilant plus occlusive are created by inserting the vowel *o*, while in feminine plural forms the gender marker /a/ avoids the difficulties caused by the presence of two sibilants in the same consonant cluster, as we see in /'ædʒ/ ‘ugly’ (*lleig*, *lletjos* versus *lletja*, *lletges*); /'trist/ ‘sad’ (*trist*, *tristos* versus *trista*, *tristes*); and /'fosk/ ‘dark’ (*fosc*, *foscós* versus *fosca*, *fosques*). We noted above a group of adjectives ending in *-ç* ([s]) that are gender-invariable in singular, such as /fe'lis/ ‘happy’ (*feliç* both in masculine and feminine singular). In this case, however, the support vowel for the feminine plural is the usual support vowel *e*,

⁵ But in Northern Catalan this *n* is also lost in plurals. For instance, the paradigm of the noun *camí* ‘way’ (< CAMMĪNU), whose general plural is *camins* ‘ways’, has become [kə'mi], [kə'mis] in this dialect.

⁶ In Old Catalan, masculine nouns ending in *-ig*, like *desig* ‘wish’, or in a sibilant followed by an occlusive, like *bosc* ‘forest’ and *gust* ‘taste’, form the plural only through the addition of the plural marker, as we see in *desigs* ‘wishes’, *boscs* ‘forests’ and *gusts* ‘tastes’. These traditional forms are maintained in formal written Catalan and can even be heard in some Balearic and Valencian varieties. The same can be said about the masculine plural inflected forms of adjectives like *lleig* ‘ugly’, *fosc* ‘dark’ and *trist* ‘sad’: *lleigs*, *foscs*, *trists*.

which in addition coincides with the manifestation of the gender marker in plurals, hence *feliços*, masculine plural, and *felices*, feminine plural.

Sporadically, there exist invariable nouns without any formal distinction between the singular and plural forms. The singular form of such nouns always ends in -s, as in *un centpeus* ‘a centipede’ versus *dos centpeus* ‘two centipedes’ or *dilluns* ‘Monday’ versus *els dilluns* ‘Mondays’. There is also a set of nouns whose singular and plural forms are orally indistinguishable but orthographically different, because the phonetic realizations of -ç and -çs are both [s], and -x and -xs are both pronounced [ks], as in *la falç* ‘the sickle’ versus *les falçs* ‘the sickles’ and *un apèndix* ‘one appendix’ versus *dos apèndixs* ‘two appendices’.

Finally, there exists a tiny group of defective nouns which only appear in plural form, although they indicate a single entity, as in *ulleres* ‘[a pair of] glasses’ and *calçotets* ‘[a pair of] underpants’.

7 Some closed classes of nominals

Having looked at the inflectional markers in nominal elements, we will now provide a brief overview of some closed classes of nominals. As we have been doing throughout the chapter, we will describe them in accordance with the most recent Catalan grammars (GCC, GIEC) with attention duly paid to etymology and dialectal diversity when appropriate (DCVB).

7.1 Articles

The Catalan definite article comes from the Latin demonstrative ILLUM, ILLAM, ILLOS, ILLAS. Its root in Catalan is /l/ and its inflected forms are *el* and *els* for the masculine, with a support vowel, and *la* and *les*, for the feminine, as in *el carrer* ‘the street’, *els carrers* ‘the streets’, *la carretera* ‘the road’ and *les carreteres* ‘the roads’. There exist apostrophised singular forms used before vowels, as in *l’ull* ‘the eye’, *l’home* ‘the man’, *l’aigua* ‘the water’ and *l’herba* ‘the grass’. In spoken Northwestern Catalan and Alguerese the older masculine forms *lo* and *los* are still used, as in *lo gos* ‘the dog’ and *los gossos* ‘the dogs’. In spoken Balearic the definite article comes from the Latin emphatic demonstrative IPSUM, IPSAM, IPSOS, IPSAS. Its Catalan root in this case is /s/ and its inflected forms are *es* for both masculine singular and plural, with a support vowel, and *sa* and *ses* for the feminine, as in *es carrer* ‘the street’, *es carrers* ‘the streets’, *sa carretera* ‘the road’ and *ses carreteres* ‘the roads’. There also exist apostrophized singular forms used before vowels, as in *s’ull* ‘the eye’, *s’home* ‘the man’, *s’aigua* ‘the water’ and *s’herba* ‘the grass’, as well as the old masculine variants *so* and *sos*, today used only after the preposition *amb* ‘with’, as in *amb so cavall* ‘with the horse’ and *amb sos cavalls* ‘with the horses’.

In turn, the indefinite article comes from the Latin numeral UNUM, UNAM. Its root is /'un/ and its inflected forms are *un*, *una*, *uns*, *unes*, as in *un carrer* 'a street', *uns carrers* '(some) streets', *una carretera* 'a road' and *unes carreteres* '(some) roads'.

In Balearic there exists a specific personal article, originating from the Latin vocatives DOMINE 'lord' and DOMINA 'lady', whose root is /n/ and which has the inflected forms *en*, for the masculine, with support vowel, and *na*, for the feminine both becoming *n'* before vowels, as in *en Joan*, *na Margalida*, *n'Antoni* and *n'Aina*. In the rest of the Catalan-speaking regions the definite article is also used before personal proper names, as in *el Jordi* and *la Núria*, except in Valencia, where the personal article is not used, as in *Vicent* and *Empar*.

7.2 Demonstratives

In Catalan there exist two demonstrative systems. Most varieties of Catalan have a binary system with two degrees of distance, near versus far, be it physically or temporally. The demonstrative used for things that are near the speech act participants, both addresser and addressee, is *aquest*, *aquesta*, *aquests* and *aquestes*, exemplified in *aquest regle* 'this ruler', *aquesta goma* 'this rubber', *aquests fulls* 'these sheets' and *aquestes carpetes* 'these folders'. The demonstratives used to refer to things that are far from the speech act participants are *aquell*, *aquella*, *aquells* and *aquelles*, as in *aquell relat* 'that story', *aquella novel·la* 'that novel', *aquells contes* 'those tales' and *aquelles biografies* 'those biographies'.

By contrast, Valencian, Ribagorçan and Eivissenc have kept the Old Catalan ternary system. This system has three degrees of distance. The demonstratives *este* (*est* in Old Catalan), *esta*, *estos* and *estes* indicate proximity to the addresser, while *eixe* (*eix* in Old Catalan), *eixa*, *eixos* and *eixes* indicate proximity to the addressee, as seen in *El document que demanen és este que tinc a les mans i no eixe que portes tu* 'The document they are asking for is this [one] which I have in my hands, not that [one] which you have brought'. Finally, distance from both of the speech act participants is expressed by the demonstratives *aquell*, *aquella*, *aquells* and *aquelles*, as in *Demana un altre imprès en aquell taulell* 'Ask for another form at that counter'.

In formal language the first-degree forms are usually replaced with *aquest*, *aquesta*, *aquests* and *aquestes*, and the second-degree forms are replaced with *aqueix*, *aqueixa*, *aqueixos* and *aqueixes*. The demonstratives *este* and *eixe* are called *simple*, while the demonstratives *aquest*, *aqueix* and *aquell* are called *reinforced*. The former come from the Latin demonstratives *ĪSTU* (> Old Catalan *est*) and *ĪPSU* (> Old Catalan *eix*), while the latter come from Vulgar Latin forms reinforced with the particle **ACCU*, coming from the interjection *ECCUM* 'here it is!', placed in front of the demonstratives, as in *ACCU-ĪSTE* > *aquest*, *ACCU-ĪPSE* > *aqueix* and *ACCU-ĪLLE* > *aquell*.

The demonstratives we have seen so far are determinants, but the transversal word class of demonstratives also includes some invariable pronouns and locative

adverbs with a deictic or anaphoric value. In the binary system the demonstrative pronouns are *això* ‘this thing (near me)/that thing (near you)’ and *allò* ‘that thing (near him/her/them)’, as in *Això que dius em deixa parat* ‘What you are saying surprises me’ and *Allò que va passar ja està oblidat* ‘What happened is already forgotten’, and the locative adverbs are *aquí* ‘here (near me)/there (near you)’ and *allí/allà* ‘there (near him/her/them)’, as in *Vine aquí!* ‘Come here!’ and *No vull que hi vagis, allí/allà* ‘I don’t want you to go there’. In the ternary system the pronouns are *açò* ‘this thing (near me)’, *això* ‘that thing (near you)’ and *allò* ‘that thing (near him/her/them)’, as in *açò meu* ‘this [is] mine’, *això teu* ‘that [is] yours’ and *allò seu* ‘that [is] his/hers/theirs’, and the locative adverbs are *ací* ‘here (near me)’, *aquí* ‘there (near you)’ and *allí/allà* ‘there (near him/her/them)’, as in *Vine ací al meu costat* ‘Come here beside me’, *No et moguis d’aquí on ets* ‘Don’t move from there where you are’ and *Si aneu allí/allà, trigareu a tornar* ‘If you go there, it will take you a long time to come back’.

7.3 Possessives

Catalan has two sets of possessive words, unstressed possessives, the use of which was common in Old Catalan but is now limited to a few specific contexts, and stressed possessives, whose use continues to be widespread. Unstressed possessives are included within the determiner class of words and appear before the noun they determine. They are clitic, that is, they have no word stress and form a stress unit with the noun they precede. The first person singular possessive has the unstressed inflected forms *mon*, as in *mon pare* ‘my father’, *ma*, as in *ma mare* ‘my mother’, *mos* as in *mos germans* ‘my brothers’, and *mes*, as in *mes germanes* ‘my sisters’. The second person singular possessive has the unstressed inflected forms *ton*, as in *ton oncle* ‘your_{SING} uncle’, *ta*, as in *ta tia* ‘your_{SING} aunt’, *tos*, as in *tos cosins* ‘your_{SING} cousins_{MASC}’, and *tes*, as in *tes cosines* ‘your_{SING} cousins_{FEM}’. By contrast, the first and second person plural have no unstressed possessive forms. Finally, the third person – both singular and plural – has the unstressed inflected forms *son*, as in *son fill* ‘his/her/their son’, *sa*, as in *sa filla* ‘his/her/their daughter’, *sos*, as in *sos fills* ‘his/her/their sons’, and *ses*, as in *ses filles* ‘his/her/their daughters’. Nowadays unstressed possessives are only used with nouns expressing a familial relation, but even in these cases the use of stressed possessives is also possible and indeed more usual, so that, for example, one is more likely to hear *la meva mare* for ‘my mother’ than *ma mare* (and even more *els meus avis* for ‘my grandparents’ than *mos avis*, because not all nouns expressing familial relationships are equally sensitive to the use of unstressed possessives).

The stressed possessives are adjectives and have a full paradigm, with forms for all singular and plural persons. The first person singular has the inflected forms *meu*, as in *el meu ordinador* ‘my computer’, *meva* or *meua*, as in *la meva/meua tauleta* ‘my tablet’, *meus*, as in *els meus apunts* ‘my notes’ and *meves* or *meues*, as in *les meves/*

meues fotocòpies ‘my photocopies’. Similarly, the second person singular has the inflected forms *teu*, as in *el teu rellotge* ‘your_{SING} watch’, *teva* or *teua*, as in *la teva/teua polsera* ‘your_{SING} bracelet’, *teus*, as in *els teus anells* ‘your_{SING} rings’ and *teves* or *teues*, as in *les teves/teues arracades* ‘your_{SING} earrings’. The third person – both singular and plural – has the inflected forms *seu*, as in *un amic seu* ‘a friend_{MASC} of his/hers/theirs’, *seva* or *seua*, as in *una amiga seva/seua* ‘a friend_{FEM} of his/hers/theirs’, *seus*, as in *uns amics seus* ‘(some) friends_{MASC} of his/hers/theirs’, and *seves* or *seues*, as in *unes amigues seves/seues* ‘(some) friends_{FEM} of his/hers/theirs’. The feminine forms with *-u-* ([w]), like *meua*, are used especially in Western Catalan and originated by analogy with the masculine forms (MEUM > *meu* → *meu* + *a* = *meua*). In Eastern Catalan, especially in Central Catalan, the *-u-* in *meua* ‘my/mine’ has become *-v-* ([v] or [β]) by analogy with other nominal elements with the alternation between *-u* in masculine (in word-final position) and *-v-* in feminine (followed by a vowel) like *blau*, *blava* ‘blue’ and *tou*, *tova* ‘soft’. Only Alguerese has conserved the Old Catalan etymological feminine forms *mia* (< MEAM) ‘my/mine’, *tua* (< TUAM) ‘your/yours_{SING}’ and *sua* (< SUAM) ‘his/her/hers’. In formal written Catalan a possessive specific to the third person plural is used whose inflected forms are *llur*, in singular, and *llurs* in plural, as in *llur fill* ‘their son’, *llur filla* ‘their daughter’, *llurs fills* ‘their sons’ and *llurs filles* ‘their daughters’. At present, only Northern varieties continue to use this possessive in the spoken language (↗8.1 Dialects).

Finally, the first person plural has the inflected forms *nostre*, as in *el nostre gos* ‘our dog’, *nostra*, as in *casa nostra* ‘our house’, and *nostres*, both masculine and feminine, as in *els nostres veïns* ‘our neighbors’ and *coses nostres* ‘our things’. Similarly, the second person plural has the inflected forms *vostre*, as in *el vostre jardí* ‘your_{PL} garden’, *vostra*, as in *la vostra piscina* ‘your_{PL} swimming pool’, and *vostres*, both masculine and feminine, as in *els vostres arbres* ‘your_{PL} trees’ and *les vostres plantes* ‘your_{PL} plants’.

7.4 Personal pronouns

We must distinguish between the stressed forms of the personal pronouns, with a word-accent, and the unstressed forms of these same pronouns, which are clitic and constitute an accentual unit when combined with the verb they precede or follow. Stressed pronouns are commonly referred to as *strong pronouns* and unstressed ones as *weak pronouns*.

The forms of the strong pronouns are as follows. The first person singular pronoun is the only one which has forms marked for case, the nominative *jo* ‘I’, used for subjects as in *Jo no vindré* ‘I do not come’, and the oblique *mi* ‘me’, used after prepositions as in *És per a mi?* ‘Is it for me?’. The second person singular pronoun is *tu* ‘you_{SING}’, as in *Tu no surts a la foto?* ‘Are you_{SING} not in the picture?’ and *Justament ara parlàvem de tu* ‘We were just talking about you_{SING}’. The first and second person

plural forms are respectively *nosaltres* ‘we/us’ and *vosaltres* ‘you_{PL}’, as in *Nosaltres ja hi som. Quan arribareu vosaltres?* ‘We are already there. When will you_{PL} arrive?’ and *No us buscaven pas a vosaltres, ens buscaven a nosaltres* ‘They weren’t looking for you_{PL}, they were looking for us’. The third person pronoun has the inflected forms *ell* ‘he’, *ella* ‘she’, *ells* ‘they_{MASC}’ and *elles* ‘they_{FEM}’ and it is the only personal pronoun for which the singular and plural forms have the same root. It is also the only strong pronoun with formal differences by gender.

The forms of the weak pronouns are as follows. The first person singular pronoun is *em* ‘me’ and it can be accusative, as in *Em criden* ‘They call me’, dative, as in *Em van donar un ram de flors* ‘They gave me a bouquet of flowers’, or reflexive, as in *Em rento* ‘I wash myself’ and *Em rento la cara* ‘I wash my face’. Similarly, the second singular pronoun is *et* ‘you_{SING}’ and it can be accusative, as in *Et busquen* ‘They are looking for you_{SING}’, dative, as in *Et compraré un llibre* ‘I will buy you_{SING} a book’, or reflexive, as in *Et canses?* ‘Are you_{SING} getting tired?’ and *Et tallaràs els cabells?* ‘Will you_{SING} cut your_{SING} hair?’. The first person plural pronoun is *ens* ‘us’, which can function as accusative, as in *Ens esperen* ‘They are waiting for us’, dative, as in *Ens ho acaben de dir* ‘They just told [it to] us’, or reflexive or reciprocal, as in *Ens veurem demà* ‘We will see each other tomorrow’. The second plural pronoun is *us* ‘you_{PL}’ and it can be accusative, as in *Us trobarem a faltar* ‘We will miss you_{PL}’, dative, as in *Us hem portat una sorpresa* ‘We have brought you_{PL} a surprise’, or reflexive or reciprocal, as in *Us vestiu?* ‘Are you_{PL} getting dressed?’. The singular and plural third person pronouns have a common root and different inflected forms depending on case, number and gender. The accusative forms are *el* ‘him/it’, *la* ‘her/it’, *els* ‘them_{MASC}’ and *les* ‘them_{FEM}’, as in *Que has vist el Joan? Fa estona que el busco* ‘Have you seen Joan? I’ve been looking for him for a while’, *Que tens el rebut? No el trobo* ‘Do you have the receipt? I can’t find it’, *Que hi ha la Maria? La demanen* ‘Is Mary here? Someone is asking for her’, *La beguda per a la festa, ja la porto jo* ‘The beverage for the party, I will bring it’, *Els convidaré tots* ‘I will invite all of them_{MASC}’, *Les convidaré totes* ‘I will invite all of them_{FEM}’. By contrast, dative forms are *li* ‘him/her’ in both masculine and feminine singular, and *els* ‘them’ in both masculine and feminine plural, as in *A la Mercè li he comprat un conte i al Roger li porto un quadern de pintar* ‘As regards Mercè, I bought her a story, and as for Roger, I’m bringing him a colouring book’, *Al Jordi i al Roger els durem una equipació de futbol del Barça* ‘As for Jordi and Roger, we will bring them an F. C. Barcelona football kit’ and *A la Núria i a la Mercè els comprarem un vestit de ballarina* ‘As for Núria and Mercè, we will buy them a dancer’s dress’. In reflexive and reciprocal uses, the third person pronoun has the form *es*, as in *Tots quatre germans es fan costat* ‘The four siblings support each other’.

Most weak personal pronouns can change their form depending on whether they occur before or after the verb, and then depending on whether the following verb begins with a vowel or consonant or the preceding verb ends in a vowel or in a consonant or glide. Some weak pronouns have four contextual forms. Thus, *em* alternates with *m’*, *’m* and *me*, as in *em molesta* ‘he/she/it bothers me’, *m’amoïna* ‘he/

she/it worries me', *avisa'm* 'let me know' and *he d'anar-me'n* 'I have to go'. Similarly, *et* alternates with *t*', *t* and *te*, as in *et miren* 'they watch you_{SING}', *t'escolten* 'they listen to you_{SING}', *atura't!* 'Stop!' and *vols quedar-te?* 'Do you_{SING} want to stay?'. This is also true of the weak pronoun *es*, which can be *es*, *s*', *'s* or *se*, as in *es perdrà* 'he/she will get lost', *s'aclareix* 'it is clearing up', *esperi's aquí* 'wait here' and *vol disfressar-se* 'he/she wants to dress up'. *El* likewise has four contextual forms, *el*, *l*', *'l* and *lo*, as in *el llegeixo* 'I read it', *l'escric* 'I write it', *repassa'l* 'revise it' and *corregeix-lo* 'correct it'.

Another set of weak pronouns has three contextual forms. For example, *ens* exists as *ens*, *'ns* and *nos*, as in *ens criden* 'they call us', *ens ignoren* 'they ignore us', *obre'ns la porta* 'open the door for us' and *vam trobar-nos* 'we met each other'. For *els* we find *els*, *'ls* and *los*, as in *els comprenc* 'I understand them_{MASC}', *els enyoro* 'I miss them_{MASC}', *cuida'ls* 'take care of them_{MASC}' and *cal protegir-los* 'it is necessary to protect them_{MASC}'. By contrast, *la* and *us* have only two forms. *La* becomes *l*' before a verb beginning with a vowel, as in *l'he vista* 'I saw her' versus *la veig* 'I see her', *vull veure-la* 'I want to see her' and *mireu-la!* 'look at her!'. *Us* alternates with *vos*, as we see in *us telefonaré* 'I will phone you_{PL}', *us admiro* 'I admire you_{PL}' and *em convé veure-us* 'I need to see you_{PL}' versus *no puc esperar-vos més* 'I can't wait for you_{PL} any longer'. Finally, a few weak personal pronouns have only one form. This is the case of *li*, as in *li fa mal* 'it hurts him/her', *li agrada* 'he/she likes him/her/it', *canta-li* 'sing to him/her' and *digueu-li la veritat* 'tell him/her the truth', and *les*, as in *les busco* 'I'm searching for them_{FEM}', *les estimo* 'I love them_{FEM}', *torna-les* 'return them_{FEM}' and *convé guiar-les* 'it is a good idea to guide them_{FEM}'.

Note that these various contextual forms can be grouped into four classes labelled *full*, *elided*, *reduced* and *reinforced*. The ten full forms are the closest to the Latin forms: *me* (< ME), *nos* (< NOS), *te* (< TE), *vos* (< VOS), *lo* (< ÍLLUM), *la* (< ÍLLAM), *los* (< ÍLLOS), *les* (< ÍLLAS), *li* (< ÍLLĪ) and *se* (< SE). The four elided forms *m*', *t*', *l*' (coming from both *lo* and *la*) and *s*' developed when a following verb began with a vowel and the pronoun could therefore form a syllable with it. Similarly, the seven reduced forms *'m*, *'ns*, *'t*, *us*, *'l*, *'ls* and *'s* came about when the pronoun followed a verb (or other word) ending in a vowel and the pronoun could form a syllable with it. Finally, the six reinforced forms *em*, *ens*, *et*, *el*, *els* and *es* were created in Modern Catalan by adding a support vowel to the reduced forms. For example, historically *te criden* 'they are calling you_{SING}' coexisted with *diu que-t criden* 'he/she says that they are calling you_{SING}', but speakers reinterpreted sequences like the latter as *diu qu-et criden*, which made it possible for *et criden* 'they call you_{SING}' to replace *te criden*. In formal written Catalan the full forms of the weak pronouns are no longer used before verbs, but they are still present in spoken Catalan, with variation among dialects. *Me*, *te* and *se* persist in Western Catalan, in some Balearic and Northern varieties and in Alguerese. *Lo* and *los* remain in Northwestern Catalan and Alguerese. *Vos* is used in Valencian, Balearic and Alguerese. *Nos*, which has become [mos] by analogy with the first person singular pronoun *me*, continues to thrive in Western Catalan, Balearic and Alguerese.

8 Inflectional markers in verbs

Verbal paradigms are much more complex than nominal ones. Each verb has 50 simple forms, the first 42 comprising the six person forms of the seven tenses, namely present indicative, as in *canta* ‘he/she sings/is singing’, simple past indicative, as in *cantà* ‘he/she sang’, imperfect indicative, as in *cantava* ‘he/she sang/was singing’, future, as in *cantarà* ‘he/she will sing’, conditional, as in *cantaria* ‘he/she would sing’, present subjunctive, as in (*volen que ell*) *canti* ‘(they want) him to sing’, imperfect subjunctive, as in (*volien que ell*) *cantés* ‘(they wanted) him to sing’. The remaining eight simple forms are two imperative forms, as in *canta!* ‘sing (you_{SING})!’ and *canteu!* ‘sing (you_{PL})!’, four participial forms, as in *cantat* ‘sung_{MASC-SING}’, *cantada* ‘sung_{FEM-SING}’, *cantats* ‘sung_{MASC-PL}’ and *cantades* ‘sung_{FEM-PL}’, the infinitive, as in *cantar* ‘to sing’, and the gerund, as in *cantant* ‘singing’.

In addition, each verb also has 68 compound forms constructed either in combination with the verb-like form *va* used as a past auxiliary to yield what is known as the *periphrastic past indicative*, as in *va cantar* ‘he/she sang’, which is fully equivalent to the simple past indicative *cantà*, or in combination with the perfect auxiliary verb to form the present perfect indicative, as in *ha cantat* ‘sang’, and other less frequently used compound tenses.⁷

With regard to the inflectional markers of Catalan verbs, we will restrict our discussion here to the regular verbs, since a full discussion of the morphology of the many irregular verbs in Catalan would carry us well outside the functional scope of this chapter.

The personal simple verb forms in Catalan can be segmented into four components that always follow the same order: the stem, the conjugation or verb class marker, the tense marker (which includes aspect and mood) and the person marker (which also includes number), as illustrated in Figure 2.

⁷ The past indicative, whether in the simple forms – used above all in formal written language – or the periphrastic forms – common in spoken language in most dialects –, is used to refer to a past action accomplished in a closed time period, as in *Ahir/La setmana passada/Fa dos anys cantà/va cantar* ‘Yesterday/last week/two years ago he/she sang’. By contrast, the present perfect indicative is used to refer to an action accomplished in a time period perceived as not closed, that is, a time period which includes the speech act, as in *Fa una estona/Avui/Aquest matí/Aquesta setmana/Aquests darrers dies ha cantat* ‘A moment ago/Today/This morning/This week/These few last days he/she has sung’.

inflected word	stem	class marker	tense marker	person marker
<i>cantàvem</i> 1 st person plural imperfect indicative, 1 st conjugation or verb class	/ˈkant/	/ˈa/	/va/	/m/
<i>perdríes</i> 2 nd person singular conditional, 2 nd conjugation or verb class	/ˈpɛrd/	∅	/ˈria/	/z/
<i>dormíreu</i> 2 nd person plural future, 3 rd pure conjugation or verb class	/ˈdɔrm/	/i/	/ˈre/	/w/
<i>parteix</i> 3 rd person singular present indicative, 3 rd inchoative conjugation or verb class	/ˈpart/	/ˈɛʃ/	∅	∅

Figure 2: Morphological segmentation of some inflected forms of the verbs *cantar* ‘to sing’, *perdre* ‘to lose’, *dormir* ‘to sleep’ and *partir* ‘to divide’.

8.1 Inflectional conjugation markers

As we have mentioned, in Catalan there exist three verb classes, or conjugations, each of them following a given inflectional pattern. Although they all share the marker /r/, it is the differences between the infinitive forms of verbs that serve as the basis for this three-way division. Infinitives of the so-called *first conjugation verbs* end in stressed *-ar*, like *cantar* ‘to sing’, and infinitives of third conjugation verbs, in a similar fashion, end in stressed *-ir*, like *dormir* ‘to sleep’. By contrast, the infinitives of the second conjugation verbs can have various endings: the most common is *-re* (unstressed), as in *perdre* ‘to lose’, but *-er* (unstressed), like *témer* ‘to fear’, *-er* (stressed), like *poder* ‘to be able to’, and *-r*, like *dur* ‘to carry’, are also possible.

The conjugation morpheme is located immediately after the stem, before the other inflectional markers, and it is usually – but not always – realised as a stressed vowel, labelled the *thematic vowel*. When the thematic vowel is not present, its place is occupied by a zero morph, as we see in *dorm* ‘he/she sleeps’ versus *dormim* ‘we sleep’, where *-i-* is the thematic vowel. Each verb class tends to be associated with a particular thematic vowel, which appears in most forms and in particular before the gerund marker /nt/. Thus, the thematic vowel of first conjugation verbs is /a/, as in *cantant* ‘singing’, the thematic vowel of second conjugation verbs is /e/, as in *perdent* ‘losing’, and the thematic vowel of third conjugation verbs is /i/, as in *dormint* ‘sleeping’. However, only in the third verb class do forms consistently exhibit the thematic vowel. In the first and second verb classes, conjugational markers undergo allomorphic variation, as can be seen in *cantàvem* ‘we sang’

versus *cantem* ‘we sing’ and in *perdem* ‘we lose’ versus *perdiem* ‘we lost’ and *perdut* ‘lost’.⁸

It should be noted that the orthographic vowel *e* can correspond to different phonological vowels. It is /*e*/ in all dialects when it appears in a gerund, as in *perdent* ‘losing’, in the simple past, as in *perderen* ‘they lost’, and in the imperfect subjunctive, as in (*volien que*) *perdés/perdera* ‘(they wanted) him/her to lose’, but it is /*e*/ in Western Catalan and Alguerese, /*ə*/ in some Balearic varieties, /*ɛ*/ in Central Catalan and other Balearic varieties and /*ɛ̃*/ in Northern Catalan when it appears in persons 4 and 5 of both the present indicative and subjunctive, as in *perdem* ‘we lose’ or *perdeu* ‘you_{PL} lose’, and the person 5 imperative, as in *perdeu!* ‘lose (you_{PL})!’.

We have also mentioned (see 3.2.1) that within the third verb conjugation there are two subclasses, the pure third class, which never receives a root extension, and the inchoative third class, with a root extension which historically comes from the Latin inchoative infix -*ESC*-. This root extension, which is always stressed, appears in persons 1, 2, 3 and 6 of the present tenses, both indicative and subjunctive, and in the person 2 imperative, which are in fact the only forms that lack the thematic vowel. In some dialects the inchoative extension has a single invariable form. This is the case of Central and Northern Catalan and Alguerese, which share the marker -*eix*-, and Northwestern Catalan, which exhibits -*ix*- as a root extension. In other dialects this inflectional marker displays allomorphy, as in Balearic between -*esc*- and -*eix*- and in Valencian between -*isc*- and -*ix*-. In these cases, the markers ending in a velar, either -*esc*- or -*isc*-, are used in the first person singular of the present indicative and in persons 1, 2, 3 and 6 of the present subjunctive, while the markers ending in a palatal are used in persons 2, 3 and 6 of the present indicative and person 2 of the imperative, as in *partesc* ‘I divide’ versus *parteix* ‘he/she/it divides’ and (*vol que*) *partisques* ‘(he/she wants) you_{SING} to divide’ versus *partix!* ‘divide (you_{SING})!’. The orthographic vowel *e* of this root extension, present only in Eastern Catalan – because Western Catalan has *i* in this inchoative marker – is pronounced [ɐ] in some Balearic varieties, [ɛ] in Central Catalan and other Balearic varieties, [ɛ̃] in Northern Catalan and [e] in Alguerese, due to its Latin origin as a long *E* (78.1 Dialects).

8.2 Inflectional tense markers

The present indicative has no tense markers. For this reason, some present indicative forms coincide with the verb root. This is the case for person 3 in verbs of the second

⁸ Balearic varieties have generally preserved the etymological thematic vowel /*a*/ in first conjugation verbs in all tenses and for this reason the first conjugation verbs in this dialect also fail to show allomorphy in this conjugational marker. Valencian has kept this etymological vowel in the imperfect subjunctive tense in association with other differences in the ending: (*volien que ell*) *cantés/cantàs* (Balearic)/*cantgra* (Valencian) ‘(they wanted) him/her to sing’.

and third conjugations, as in *perd* ‘he/she loses’ and *dorm* ‘he/she sleeps’. By contrast, the present indicative forms for the persons 2 and 6 of these conjugations have two morphemes with phonological content, the root and the person marker, as in *perds* ‘you_{SING} lose’ and *dorms* ‘you_{SING} sleep’, and *perden* ‘they lose’ and *dormen* ‘they sleep’, in which a support vowel is necessary to make the syllabification of the person marker /n/ possible.⁹ Person 1 reflects dialectal variation, acting like person 3 when it does not receive a person marker, as in *perd* ‘I lose’ and *dorm* ‘I sleep’, but acting like person 2 when it receives a person marker, as in *perdo* ‘I lose’ and *dormo* ‘I sleep’. (We examine these dialectal differences in greater detail in 8.3 below.) In persons 4 and 5 of all three conjugations, the morphological segmentation of the present indicative inflected forms presents three morphs: the root, the stressed thematic vowel and the person marker, as we see in *cantem* ‘we sing’ = *cant* + *e* + *m* or *perdeu* ‘you_{PL} lose’ = *perd* + *e* + *u*. Finally, persons 2, 3 and 6 of the first conjugation show a vowel that was historically the thematic vowel of Latin first class verbs, as in *cantes* (< CANTAS) ‘you_{SING} sing’, *cantā* (< CANTAT) ‘he/she sings’ and *canten* (< CANTANT) ‘they sing’. In fact, this vowel defies easy classification. One option is to continue to regard it as a thematic vowel; this line is easy to argue on the basis of history, but it clashes with the modern rule stating that thematic vowels must be stressed. The other option would be to regard this vowel as a present tense marker, but this violates the rule according to which the present indicative tense does not exhibit tense markers due to its unmarked character.

The past indicative marker has three allomorphs distributed in the same way in all three conjugations: the zero morph in person 3, as in *cantà* ‘he/she sang’, *perdé* ‘he/she lost’ and *dormí* ‘he/she slept’, /i/ in person 1, as in *cantí* ‘I sang’, *perdí* ‘I lost’ and *dormí* ‘I slept’, and /re/ everywhere else, as in *cantaren* ‘they sang’, *perderen* ‘they lost’ and *dormiren* ‘they slept’.¹⁰

The imperfect indicative marker is /va/ in the first conjugation, but /a/ in the second and third conjugations. These tense markers are spelled -va and -a in absolute word-final position and -ve- and -e- when followed by a consonant or glide, as in (jo) *cantava* ‘I sang’ versus *cantaven* ‘they sang’ and (jo) *perdia* ‘I lost’ versus *perdien* ‘they

⁹ A support vowel is also necessary in persons 2 and 6 of the inchoative third conjugation verbs to make possible the adjunction of the person markers /z/ and /n/ after the root extension ending in a sibilant, as in *parteixes* ‘you divide’ and *parteixen* ‘they divide’.

¹⁰ Note that in the third person singular past indicative *dormí* ‘he/she slept’ the last vowel /i/ is the thematic vowel of third conjugation verbs, just as /a/ in *cantà* ‘he/she sang’ is the thematic vowel of first conjugation verbs and /e/ in *perdé* ‘he/she lost’ is the thematic vowel of second conjugation verbs, but in the first person singular past indicative *dormí* ‘I slept’ – apparently identical to the third person singular form – the final vowel /i/ is the past tense marker rather than the thematic vowel. We know this because exactly the same vowel appears in the analogous first person singular forms of the other two conjugations, as in *cantí* ‘I sang’ and *perdí* ‘I lost’.

lost'. This vocalic alternation has the same explanation as the variation in the feminine marker /a/, spelled -a in singular, *rosa*, and -e in plural, *roses*.

In all Catalan dialects the future marker presents the allomorph /'ra/ in persons 2, 3 and 6, as in *cantaràs* 'you_{SING} will sing', *perdrà* 'he/she will lose' and *dormirà* 'they will sleep'. In the first person singular the marker is usually /'re/, although in some northwestern varieties it is /'rɛ/ due to the different evolutionary path of the Romance diphthong [aj], as in *cantaré* 'I will sing', *perdré* 'I will lose' and *dormiré* 'I will sleep'. Finally, in persons 4 and 5, as in *cantarem* '(we) will sing' and *dormireu* 'you_{PL} will sleep', this tense marker is /'re/ in Western Catalan and Alguerese, /'rə/ in some Balearic varieties, /'rɛ/ in Central Catalan and other Balearic varieties, and /'rɛ/ in Northern Catalan, reflecting the variable evolution of the Latin long E in the Catalan dialects (78.1 Dialects).

The conditional marker is /'ria/ in all persons and conjugations. It is spelled -ria in absolute word-final position and -rie- when followed by a consonant or glide, as in (jo) *cantaria* 'I would sing' versus *cantarien* 'they would sing'. It should be remembered that this vowel distinction is phonetic in Western Catalan, [kanta'ria] or [kanta'rie] 'I would sing' versus [kanta'rien] 'they would sing', but neutralised in Eastern Catalan, [kəntə'riə] 'I would sing' and [kəntə'riən] 'they would sing'.

The present subjunctive has tense markers in persons 1, 2, 3 and 6, but it presents the zero morph in persons 4 and 5, which are syncretic with the same persons of the present indicative: (*volen que*) *cantem* '(they want) us to sing', (*volen que*) *perdem* '(they want) us to lose' and (*volen que*) *dormim* '(they want) us to sleep'. In contemporary Eastern Catalan, the present subjunctive marker is /i/ in all conjugations: (*volen que*) *cantis* '(they want) you_{SING} to sing', (*volen que*) *perdis* '(they want) you_{SING} to lose', (*volen que*) *dormis* '(they want) you_{SING} to sleep'. By contrast, Western Catalan usually conserves older forms with the marker /e/ in the first conjugation, as in (*volen que*) *cante* '(they want) me/him/her to sing' and (*volen que*) *cantes* '(they want) you_{SING} to sing', and with the marker /a/ in the second and third conjugations, as in (*volen que*) *perda* '(they want) me/him/her to lose' and (*volen que*) *dorma* '(they want) me/him/her to sleep'. As usual, when the marker /a/ appears in a closed syllable it becomes [e], as in (*volen que*) *perdes* '(they want) you_{SING} to lose' and (*volen que*) *dormes* '(they want) you_{SING} to sleep'.

The imperfect subjunctive has tense markers in all persons. In Eastern Catalan there exist two allomorphs: /s/, which appears in persons 1 and 3, as in (*volien que*) *cantés* '(they wanted) me/him/her to sing', (*volien que*) *perdés* '(they wanted) me/him/her to lose' and (*volien que*) *dormís* '(they wanted) me/him/her to sleep', and /si/, which appears in persons 2, 4, 5 and 6, as in (*volien que*) *cantessis* '(they wanted) you_{SING} to sing', (*volien que*) *perdessis* '(they wanted) you_{SING} to lose' and (*volien que*) *dormissis* '(they wanted) you_{SING} to sleep'. Northwestern Catalan has a very similar system, the only difference being that the second allomorph is the classical /se/, as in (*volien que*) *cantesses* '(they wanted) you_{SING} to sing', (*volien que*) *perdesses* '(they wanted) you_{SING} to lose' and (*volien que*) *dormisses* '(they wanted) you_{SING} to sleep'.

By contrast, Valencian uses a different system, derived from a separate Latin origin, which consists in the insertion of the marker /ra/ in all persons. This is pronounced [ra] in absolute final position, as in (*volien que*) *cantara* ‘(they wanted) me/him/her to sing’, (*volien que*) *perdera* ‘(they wanted) me/him/her to lose’ and (*volien que*) *dormira* ‘(they wanted) me/him/her to sleep’, but [re] in closed final syllable, as in (*volien que*) *cantares* ‘(they wanted) you_{SING} to sing’, (*volien que*) *perderes* ‘(they wanted) you_{SING} to lose’ and (*volien que*) *dormires* ‘(they wanted) you_{SING} to sleep’.

The only two persons of the imperative mood, persons 2 and 5, do not show tense markers and are syncretic with respect to persons 3 and 5 of the present indicative, as can be seen in *canta!* ‘sing (you_{SING})!’ (like *canta* ‘he/she sings’) and *dormiu!* ‘sleep (you_{PL})!’ (like *dormiu* ‘you_{PL} sleep’).

8.3 Inflectional person markers

As we have seen, in Catalan verbal inflection the grammatical properties of person and number always appear fused into a single inflectional marker. The first and third persons singular have no marker and therefore usually exhibit syncretism, as can be seen in *Jo cridava, però ell també cridava* ‘I shouted, but he also shouted’ and *No volien només que ell cantés, sinó que també pretenien que cantés jo* ‘They not only wanted him to sing, they also tried to get me to sing’.¹¹ The other persons have these inflectional markers: /z/, second person singular, as in *cantes* ‘you_{SING} sing’, *perdràs* ‘you_{SING} will lose’ and *dormies* ‘you_{SING} slept’; /m/, first person plural, as in *cantem* ‘we sing’, *perdrem* ‘we will lose’ and *dormiem* ‘we slept’; /w/, second person plural, as in *canteu* ‘you_{PL} sing’, *perdreu* ‘you_{PL} will lose’ and *dormieu* ‘you_{PL} slept’; and finally /n/, third person plural, as in *canten* ‘they sing’, *perdran* ‘they will lose’ and *dormien* ‘they slept’.

Although these generalisations with regard to inflection for person do tend to hold, there is in fact considerable variation across dialects with regard to the first person singular in the present indicative (78.1 Dialects). Balearic and Alguerese exhibit no marker at all, as in *cant* ‘I sing’, *perd* ‘I lose’ and *dorm* ‘I sleep’. Valencian has the marker -e, but only in first conjugation verbs, as in *cante* ‘I sing’. Northern Catalan or Rossellonese has the marker -i in all conjugations, as in *canti* ‘I sing’, *perdi*

¹¹ Northwestern Catalan is the only dialect that manages to avoid this syncretism of person through various means. For example, here we find that while the Latin ending -AM has evolved into -a, as in general Catalan, the Latin ending -AT has evolved into -[e]: *cantava* ‘I sing’ versus *cantav[e]* ‘he/she sings’, *cantaria* ‘I would sing’ versus *cantari[e]* ‘he/she would sing’. Similarly, in the subjunctive tenses the first person singular ending -a is contrasted with the third person singular ending in Northwestern Catalan: (*volen que*) *canta* ‘(they want) me to sing’ versus (*volen que*) *canto/cante/canti* ‘(they want) him/her to sing’ and (*volien que*) *cantessa* ‘(they wanted) me to sing’ versus (*volien que*) *cantés* ‘(they wanted) him/her to sing’.

'I lose' and *dormi* 'I sleep'. Finally, within the territory of (Southern) Catalonia the marker is -o in all conjugations, as in *canto* 'I sing', *perdo* 'I lose' and *dormo* 'I sleep', though it is pronounced [o] in Northwestern Catalan and [u] in Central Catalan, in accordance with the usual manifestation of unstressed vowels in these two dialects.

9 Bibliography

- Badia i Margarit, Antoni Maria (1981), *Gramàtica històrica catalana*, València, Tres i Quatre.
- Badia i Margarit, Antoni Maria (1994), *Gramàtica de la llengua catalana: descriptiva, normativa, diatòpica, diastràtica*, Barcelona, Enciclopèdia Catalana.
- Batlle, Mar, et al. (2016), *Gramàtica històrica de la llengua catalana*, Barcelona, Publicacions de l'Abadia de Montserrat.
- DCVB – Alcover, Antoni Maria/Moll, Francesc de Borja (1993, ¹1926–1962), *Diccionari català-valencià-balear*, 10 vol., Palma, Moll, <http://dcvb.iecat.net/>.
- Fabra, Pompeu (⁷1933, ¹1918), *Gramàtica catalana*, Barcelona, Institut d'Estudis Catalans.
- Fabra, Pompeu (1956), *Gramàtica catalana*, Barcelona, Teide.
- GCC = Solà, Joan, et al. (edd.) (²2008, ¹2002), *Gramàtica del català contemporani*, 3 vol., Barcelona, Empúries.
- GIEC = Institut d'Estudis Catalans (2016), *Gramàtica de la llengua catalana*, Barcelona, Institut d'Estudis Catalans.
- Mascaró, Joan (1986), *Morfologia catalana*, Barcelona, Enciclopèdia Catalana.
- Moll, Francesc de Borja (1991), *Gramàtica històrica catalana*, València, Universitat de València.
- Pérez Saldanya, Manuel (1998), *Del llatí al català. Morfosintaxi verbal històrica*, València, Universitat de València.
- Veny, Joan (¹²1998, ¹1978), *Els parlars catalans*, Palma de Mallorca, Moll.

5.2 The Simple Sentence

Abstract: This chapter deals with simple declarative sentences and their main properties. In section 1, some of the most salient properties of simple sentences in Catalan, as a Romance language, are schematically described. Section 2 offers an approach to sentence constituents and their syntactic functions. Section 3 is devoted to the head of the NP, the noun and its classes, and to the specifiers of the NP: articles, demonstratives, possessives, and quantifiers (numerals, quantitatives, indefinites and universals). Special attention is paid to partitive constructions in 3.2.4. In section 4, the syntactic functions of personal strong and weak pronouns are described. Other weak pronouns – *en*, *hi* and *ho* – are studied in 4.2. The types of adjectives and the constituents of the AdjP are described in section 5, while the behavior of stressed and unstressed prepositions and that of adverbs are presented in sections 6 and 7, respectively. Section 8 examines the properties of verbs and verb phrases and addresses the VP's constituents and types of verbs (transitive verbs in 8.1, intransitive in 8.2, unaccusative verbs and constructions in 8.3 and copular, quasi-copular and presentational verbs in 8.4). Section 9 is devoted to verbal tenses (the indicative and subjunctive tenses), while section 10 deals with the aspectual and modal verbal periphrases. Finally, affirmative polarity and negative polarity in declarative sentences are described in section 11.

Keywords: Catalan syntax, lexical categories, phrases, tense and aspect, verbal periphrases, polarity

1 Properties of the Catalan simple sentence

Catalan shares a lot of syntactic properties with other Romance languages. For instance, Catalan sentences – contrary to French, but similarly to Portuguese, Italian, Occitan or Spanish – can have an overt or covert subject, as shown in (1). The presence or absence of the subject depends on the conditions of the discourse or conversation (see Rigau 1986).

- (1) a. Tu cantaves.
 ‘You were singing’
 b. Cantaves.
 ‘You were singing’

The second person singular morpheme of the verbal form (-s in *cantaves*) enhances the omission of the subject. Moreover, the subject can be emphasized by placing it in post-verbal position: *Cantaves tu*.

Nevertheless, Catalan exhibits some syntactic peculiarities. Let us take a look at some of them:

a) Catalan has a strong tendency to dislocate the objects and adjuncts of the predicate (see (2a) and (2b) in relation to (2c); and (2d) and (2e) in relation to (2f)). This situation is favored especially by the weak or clitic pronominal system (see 4.1–4.2), which is richer than in other Romance languages, such as Portuguese and Spanish. The definite weak pronouns and the weak pronouns *ho*, *en* and *hi* can also be found in French, Occitan or Italian. However, their syntactic values do not coincide entirely with those of Catalan weak pronouns. On dislocation in Catalan see 75.4 Modality and Information Structure: Focus, Dislocation, Interrogative and Exclamatory Sentences.

- (2) a. De llibres, al pare, ella li'n regala tres.
of books, to-the father she to-her *partitive clitic* gives three
b. Ella li'n regala tres, de llibres, al pare.
she to-her *partitive clitic* gives three, of books, to-the father
c. Ella regala tres llibres al pare.
'She gives three books to the father'
d. El cotxe, al taller, l'hi portarem demà.
the car to the garage it *locative clitic* will take.1.PL tomorrow
e. L'hi portarem demà, el cotxe, al taller.
'it *locative clitic* will take.1.PL tomorrow, the car, to the garage'
f. Portarem el cotxe al taller demà.
'We will take the car to the garage tomorrow'

b) Similarly to other Romance languages, some verbs of movement are also used as auxiliary verbs in Catalan. However, Catalan is the only Romance language that exhibits a past tense expressed by means of forms derived from the present of the verb *anar* 'go' + infinitive. This tense is called periphrastic past, and it coexists with a simple past, as shown in (3).

- (3) a. Ahir ella *va escriure* el poema.
yesterday she *goes to write* the poem
'Yesterday she *wrote* the poem'
b. Ahir ella *escrigué* el poema.
'Yesterday she *wrote* the poem'

c) Catalan negative sentences usually contain the negative adverb *no* and, in some dialects, *no ... pas*. The use of the negative adverb *pas* in Catalan does not coincide

exactly with that of the French and Occitan *pas*. It is used mainly to negate what the listener presupposes to be true or to place emphasis on the negation (see 11).

- (4) Speaker A: – Aquesta tarda escriurem la carta.

‘This afternoon we will write the letter’

- Speaker B: – Aquesta tarda *no* vindré *pas*.

this afternoon *not* will-come.1.SG *pas*

‘This afternoon I WONT’ come’

d) Catalan, similar to other languages, exhibits a two-term space deixis: *aquí* ‘here’ and *allà* (or *allí*) ‘there’. However, the distinction between the two is as follows: *aquí* may refer to either the speaker’s location (5a) or the addressee’s location (5b), while *allà* (or *allí*) is a location distant from both the speaker and/or the addressee (5c). See 3.1b for more details.

- (5) a. Aquí on sóc jo.

‘Here where I am’

- b. Aquí on ets tu.

‘Here where you are’

- c. Allà on és ell.

‘There where he is’

e) Headless relative clauses can appear in the indicative mood, similarly to Gallo-Romance languages, or in the subjunctive, similarly to Ibero-Romance languages.

- (6) a. Aniràs a la platja quan *acabaràs* els deures.

will-go.2.SG to the beach when will-finish.2.SG your homework

‘You will go to the beach when you finish your homework’

- b. Aniràs a la platja quan *acabis* els deures.

will-go.2.SG to the beach when finish.2.SG.subj your homework

‘You will go to the beach when you finish your homework’

f) The presence of negative polarity items (*res* ‘nothing’, *ningú* ‘nobody’, *cap* ‘none’) in interrogative, conditional and dubitative sentences with a positive interpretation is another idiosyncratic property of Catalan:

- (7) a. Vols *res* del mercat?

want.2.SG nothing from the market?

‘Do you want *anything* from the market?’

- b. Si ve *ningú*, avisa’m.

if *nobody* comes, let me know

‘If *anyone* comes, let me know’

- c. Dubto que *cap* autobús vagi a l'aeroport.
 doubt.1.SG that *none* bus goes to the airport
 'I doubt that *any* bus goes to the airport'

In the following sections, the constituents of declarative simple sentences and the types of sentences according to their predicates are presented.

2 Constituents and functions

Words are hierarchically arranged within the sentence in order to form wider constituents or phrases, which are labeled according their head: noun phrase (NP), verb phrase (VP), adjective phrase (AdjP), preposition phrase (PP), and adverb phrase (AdvP). Such phrases have different grammatical functions in the sentence: prototypically, subject of the sentence, predicate of the sentence, object(s) of the predicate and adjuncts of the predicate (or VP-oriented adjuncts), among others. We assume that the canonical word order in Catalan sentences is subject-verb-object (see 75.4 Modality and Information Structure: Focus, Dislocation, Interrogative and Exclamatory Sentences). If the sentence is simple, the subject is a NP and the verb projection is a VP that contains the verb, its selected objects and the adjuncts of the predicate, if any. For more details on constituents and grammatical functions, see Hernanz (2002), Bel (2002) and GIEC (2016, 471–569).

The NP has a nominal head and may appear with specifiers and complements. The head of the NP is a noun (8a) or a pronoun (8b) and (8c):

- (8) a. [_{NP} Barcelona]
 b. [_{NP} Ella] 'she'
 c. [_{NP} Ningú] 'no one'

The NPs in (9a,b) have an article and a demonstrative as specifier, while those in (9c,d) have an AdjP and a PP as a complement (or modifier):

- (9) a. [_{NP} La casa]
 'the house'
 b. [_{NP} Aquella casa]
 'that house'
 c. [_{NP} La [casa [_{AdjP} més gran]]]
 the house more big
 'the biggest house'
 d. [_{NP} Aquella [casa [_{PP} de la plaça]]]
 'that house on (*lit.* of) the square'

The verb in (10a) selects a NP as a direct object to form the VP. The verb in (10b) selects an NP and a PP as direct object and indirect object, respectively. The verb in (10c) selects a PP as its object (or prepositional regime supplement), whereas the verb in (10d) does not select any object (see 8.2).

- (10) a. [_{VP} llegeix [_{NP} un llibre]]
 ‘reads a book’
 b. [_{VP} donarem [_{NP} la clau] [_{PP} al veí]]
 ‘will-give.1.PL the key to the neighbor’
 c. [[_{VP} parlen [_{PP} de política]]
 talk.3.PL of politics
 ‘they talk about politics’
 d. [_{VP} salta]
 ‘jumps’

Optionally, a VP can appear with an adjunct as a predicate modifier, generally in the form of a PP (e.g., *al pati* ‘in the patio’) or an AdvP (e.g., *desganadament* ‘without much interest’):

- (11) [_{VP} renta [_{NP} el cotxe] [_{PP} al pati] [_{AdvP} desganadament]]
 ‘washes the car in the patio without much interest’

Moreover, a PP or an AdvP can act as a sentence-oriented adjunct. This is the case of the PP *a València* ‘in Valencia’ and the AdvP *educadament* ‘politely’ in (12), which appear in the left periphery of the sentence (cf. Rigau 2002, 2054–2057).

- (12) a. A València, em llevo d’hora.
 ‘In Valencia, I wake up early’
 b. Educadament, ella no va replicar.
 ‘Politely, she didn’t reply’

The PP *a València* in (12a) can be paraphrased as “when I am in Valencia”, and (12b) can be paraphrased as “It was polite of her not to reply”.

Furthermore, other elements such as conjunctions have to be considered. Coordinating conjunctions must be distinguished from subordinating ones. Coordinating conjunctions may be copulative (*i* ‘and’, *ni* ‘nor’), adversative (*però*, *sinó* ‘but’), disjunctive (*o*, *o bé* ‘or’), among others. The main subordinating conjunctions are *que* ‘that’, *si* ‘whether, if’, *perquè* ‘because’, and compound conjunctions such as *ja que* ‘since’, *com que* ‘as’, *encara que* ‘although’. As for coordinating and subordinating conjunctions, see 7.5.3 The Complex Sentence.

- (13) a. l'home i la dona
 'the man *and* the woman'
 b. l'home o la dona
 'the man *or* the woman'
 c. No són cosins *sinó* germans
 'They are not cousins *but* brothers'

3 Nouns and noun phrases

Catalan nouns inherently exhibit grammatical gender, and most of them can be inflected for number (for other variants see 75.1 Word Classes, Inflectional Categories and Paradigms). They can be feminine or masculine, and adjectives and determiners agree with nouns in gender and number: *cadira* (f.) 'chair', *llit* (m.) 'bed', *mestre* (m.) 'teacher', *mestra* (f.) 'teacher'. However, sometimes the same word is used for the masculine and the feminine: *el lingüista francès* (m.) 'the French linguist', *la lingüista francesa* (f.) 'the French linguist'.

Catalan nouns can be proper nouns or common nouns. There are various types of proper nouns: personal names (*Maria, Joan, Pep*) and surnames (*Fabra, Moll, Sanchis*), geographical names (*Catalunya, Maó, Mediterrani*), and also the titles of books or plays (*Antígona*), buildings (*Palau de les Arts*), hurricanes (*Otto*), etc. Among common nouns, it is useful to distinguish between countable nouns, used in singular or plural (*taula* 'table', *gat* 'cat', *ciutat* 'town') and uncountable ones, only used in singular (*aigua* 'water', *sal* 'salt', *sinceritat* 'sincerity'); and between concrete nouns (*casa* 'house', *pa* 'bread', *sabó* 'soap') and abstract nouns (*veritat* 'truth', *prudència* 'prudence', *pau* 'peace'). Some common nouns have a collective meaning (*grup* 'group', *família* 'family', *societat* 'society') while others have a quantitative meaning (*munt* 'pile', *litre* 'litre', *grapat* 'handful'). Finally, some infinitive forms are used as nouns (*el despertar de la sexualitat* 'the awakening of the sexuality'). On nominalizations see Martí i Girbau (2002, 1328–1332), GIEC (2016, 529–535).

3.1 Determiners

Nouns may appear with a determiner, as shown above in example (9). There are several types of determiners: articles, demonstratives, possessives, and nominal quantifiers are the main ones.

a) Articles mark gender and number in agreement with the noun they specify. The definite articles are *el* or *l'* (m. sg.), *la* or *l'* (f. sg.), *els* (m. sg.), *les* (f. pl.): *el mar* 'the sea' or *l'home* 'the man', *la sala* 'the lounge' or *l'ànima* 'the soul', *els núvols* 'the clouds', *les estrelles* 'the stars'. The indefinite articles are *un* (m. sing.), *una* (f. sg.), *uns* (m. sg.), *unes* (f. sg.): *un home* 'a man', *una taula* 'a table', *uns ocells* 'some birds', *unes*

mntanyes ‘some mountains’. Depending on the dialect, personal proper nouns can appear with the definite article (*el*, *la*) or a specific personal article (*en*, *na* from vocative Latin DOMINE ‘man’ and DOMINA ‘woman’, respectively): *El Pere* (l’*Antoni*) or *en Pere* (n’*Antoni*), *la Maria* (l’*Antònia*) or *na Maria* (n’*Antònia*). This personal article is dropped in Valencian (*Pere*, *Maria*). These articles are not present in vocative forms: *Oh, Maria!* On the other hand, some geographical names contain the definite article: river names (*el Ter*, *la Muga*), the majority of mountains (*el Montseny* vs. *Montserrat*), and the names of certain towns, cities and villages (l’*Hospitalet de l’Infant*, l’*Havana*, vs. *Barcelona*, *València*).

Moreover, Catalan has the individualizer or neuter article *el* or *l’* (colloquially *lo*), which individualizes a property. It can appear with an AdjP or in a free relative clause. In this case, *el* can be substituted by the distal demonstrative pronoun *allò* ‘that’: *el més normal* ‘what is most usual’, *l’important* ‘what is important’, *el que cal fer* (or *allò que cal fer*) ‘what we must do’, etc.

b) Like demonstrative adverbs, demonstrative determiners exhibit a two-term space deixis with the following forms: *aquest* (m. sg.)/*aquesta* (f. sg.)/*aquests* (m. pl.)/*aquestes* (f. pl.) ‘this/these’ to point at something close to the speaker and/or the addressee, and *aquell* (m. sg.)/*aquella* (f. sg.)/*aquells* (m. pl.)/*aquelles* (f. pl.) ‘that, those’ to refer to something distant from both the speaker and the addressee (see § 1c): *aquesta casa* ‘this house’, *aquells arbres* ‘those trees’. In some dialects, such as Valencian, there is a form for expressing mediate proximity or proximity to the addressor: *aqueix/aqueixa/aqueixos/aqueixes*: *aqueixes sabates teues* ‘those shoes of yours’ (for other variants see 75.1 Word Classes, Inflectional Categories and Paradigms). Demonstratives can appear with an elliptical noun retrievable from the context of the sentence of discourse: *He rentat aquell plat i aquest* (meaning *aquest plat*) ‘I washed that dish and this one’. In such uses, *aquest* has been considered a demonstrative pronoun by traditional grammarians (see Brucart 2002, 1438–1445). However, the undisputable demonstrative pronouns are *açò* ‘this’, *això* ‘this’, and *allò* ‘that’. These cannot co-occur with a noun. On demonstrative adverbs, see 7d.

c) Catalan has two types of possessive elements: weak (or unstressed) possessives and strong possessives. Both express grammatical person and agree with the noun in gender and number. Weak possessives are defective (lacking first and second plural persons):

Table 1: Weak possessives.

PERSON	SINGULAR		PLURAL	
	MASCULINE	FEMININE	MASCULINE	FEMININE
1 st sg.	mon	ma	mos	mes
2 nd sg.	ton	ta	tos	tes
3 rd sg. or pl.	son	sa	sos	ses

These forms are restricted to nouns denoting familial relationships or other nouns closely related to personal intimacy, mainly *casa* ‘home’ and *vida* ‘life’: *mon pare* ‘my father’, *ta mare* ‘your mother’, *ses ties* ‘his/her/their aunts’, *ma casa* ‘my home’, *en ma vida* (‘in my life’, meaning ‘never’).

Strong possessives are more frequently used and in general they appear in the following forms (for other variants see 75.1 Word Classes, Inflectional Categories and Paradigms).

Table 2: Strong possessives.

PERSON	SINGULAR		PLURAL	
	MASCULINE	FEMININE	MASCULINE	FEMININE
1 st singular	meu	meva	meus	meves
2 nd singular	teu	teva	teus	teves
3 rd singular	seu	seva	seus	seves
1 st plural	nostre	nostra		nostres
2 nd plural	vostre	vostra		vostres
3 rd plural	seu	seva	seus	seves

Unlike weak possessives, strong ones do not occupy a determiner position, but rather must be preceded by a determiner (*el meu pare* ‘my father’) in prenominal position. In fact, when combined with the definite article or a demonstrative, the possessive can be prenominal (*el teu llibre*, *aquella seva filla*, literally ‘the your book, that his/hers/their daughter’) or, less frequently, postnominal (*el llibre teu*, *aquella filla teva*, literally ‘the book of yours, that daughter of yours’). On the other hand, strong possessives occur in postnominal position when combined with the indefinite article or a nominal quantifier in prenominal position: *un llibre meu* ‘a book of mine’, *uns amics teus* ‘some friends of yours’, *tres amigues seves* ‘three friends of his/hers/theirs’, *molts llibres nostres* ‘a lot of our books’, *algun amic vostre* ‘some friend of yours’. Finally, the strong possessive is always postnominal when the noun phrase has no specifier, as in *És cosa meva* (literally ‘it is thing mine’), and in a vocative construction: *Filla meva!*, literally ‘daughter mine’.

3.2 Nominal quantifiers

Nominal quantifiers are numerals, quantitatives, indefinites and universals. On nominal quantifiers with an interrogative or exclamative meaning (*quin producte?* ‘which product’, *quantes roses?* ‘how many roses?’, *quina calor!* ‘what a heat!’), see 75.3 The Complex Sentence.

3.2.1 Numeral quantifiers

Numeral quantifiers include cardinals, fractions, collective numerals, and multiple numerals. Ordinal numerals (*primer* ‘first’, *trentè* ‘thirtieth’, etc.) are not quantifiers, but adjectives based on cardinal numerals (see 5d).

a) Cardinal numerals such as *set* ‘seven’, *tres-cents* ‘three hundred’, etc. can appear preceded by a demonstrative or a definite article: *aquelles vint cases* ‘those twenty houses’, *les cent preguntes* ‘the one hundred questions’. *Un* ‘one’ and *dos* ‘two’ show gender inflection (*una* and *dues*): *dues-centes ovelles* ‘two hundred sheep’. When preceding a cardinal numeral, the indefinite article *uns/unes* means ‘approximately, about’: *uns tres mil euros* ‘about three thousand euros’. On the contrast between cardinal numbers in Catalan and English, see Wheeler/Yates/Dols (1999, 150–155). In addition to their quantifying role, cardinals are used as nouns to denote natural numbers and numbered entities: *Cinc i u són nombres senars* ‘Five and one are odd numbers’, *l’autobús 14* ‘route 14’.

b) Fractions are numeral quantifiers that denote a part of a unit: *mig/mitja* ‘half’, *meitat* ‘half’, *dècim* ‘tenth’, *terç* ‘third’, *milionèsim* ‘1,000,000th’, etc.:

- (14) a. *mig meló/mitja poma*
 ‘half a melon/half an apple’
 b. *la meitat d’una poma*
 ‘half an apple’
 c. *un dècim dels teus guanys*
 ‘a tenth of your income’
 d. *dos quarts dels vots emesos*
 ‘two quarters of the votes cast’

Mig and *meitat* express the fraction ‘half’. *Mig* is the only fraction quantifier that acts as a specifier and agrees in gender with the noun, which is countable and singular, as shown in (14a). It can be coordinated with a NP headed by a cardinal numeral (*dues setmanes i mitja* ‘two and a half weeks’). Other fractions co-occur with the definite article (14b), the indefinite article (14c) or a cardinal numeral (14d) and the preposition *de* ‘of’.

c) Collective numeral quantifiers are mainly formed by combining a collective noun (*parell* ‘pair, couple’, *dotzena* ‘dozen’, *vintena* ‘score’, *centenar* ‘hundred’, etc.) and a previous cardinal numeral. The preposition *de* ‘of’ links the collective quantifier with the specified noun: *una dotzena d’ous* ‘a dozen eggs’, *dos parells de sabates* ‘two pairs of shoes’, *tres centenars de llibres* ‘three hundred books’.

d) Multiple numeral quantifiers express the number of times that an entity increases: *doble* ‘double’, *triple* ‘triple’, *quàdruple* ‘quadruple’, *quíntuple* ‘quintuple’, etc. They can be used as quantifiers (*Pagaran doble salari* ‘A double salary will be paid’) or as nouns (*Sis és el doble de tres* ‘Six is the double of three’).

3.2.2 Quantitative quantifiers

Quantitative quantifiers indicate quantity or amount depending on whether the noun they specify is countable or uncountable. Some of them are inflected in gender and number: *molt (de)* ‘much, many’ (*molts (de) llibres* ‘many books’); *bastant (de)* ‘enough’ (*bastanta aigua* ‘enough water’), *poc* ‘a little, few’ (*poques flors* ‘few flowers’), etc. Others are invariable: *força llibres* ‘a lot of books’; *gens de* ‘no’ (*gens d’ajuda* ‘no help at all’); *massa* ‘too many/much’ (*massa cadires* ‘too many chairs’); *més* ‘more’; *menys* ‘less’; *prou* ‘enough’; etc. The quantitative quantifiers *gaire* ‘not many/much’ and *gens de* ‘nothing of, no, any’ occur in negative constructions (*No tinc gaires parents* ‘I don’t have many relatives’; *No tinc gens de dolor* ‘I have no pain’) or in interrogative, conditional and dubitative constructions: *Tens gens de dolor?* ‘Do you have any pain?’. On degree (or quantitative) adverbs related to quantitative nominal quantifiers (*molt, poc, prou, gens*, etc.), see 7h.

Other quantitative quantifiers include measure nouns such as *litre* ‘litre’, *pam* ‘handspan’, *gram* ‘gram’, etc., and other nouns such as *munt* ‘pile’, *grapat* ‘handful’, etc.: *cent grams de mantega* ‘one hundred grams of butter’, *un grapat d’arròs* ‘a handful of rice’. These phrases are called pseudopartitive (see Brucart/Rigau 2002, 1538–1545; GIEC 2016, 642–643).

3.2.3 Indefinite and universal quantifiers

a) Indefinite quantifiers usually express existence (*algun* ‘some’, *cert* ‘one, certain’, *altre* ‘other’, *uns quants* ‘some’, *qualsevol* ‘any’ or absence of existence (*cap* ‘no’): *alguna casa* ‘some house’, *cert home* ‘a certain man’, *unes quantes coses* ‘some things’, *qualsevol llibre* ‘any book’, *cap cas* ‘no case’. Some indefinite words are pronouns: *algú* ‘somebody’, *qualcom* ‘something’, and the negative pronouns *res* ‘nothing’ and *ningú* ‘nobody’.

b) Universal quantifiers apply to all the entities of a set (*tot* ‘every, all’, the distributive quantifier *cada* ‘each’, *ambdós* ‘both’, *sengles* ‘each’): *tot llibre* ‘every book’, *cada estudiant* ‘each student’, *ambdós amics* ‘both friends’. *Tot* can quantify a definite NP, acting as a predeterminer: *tots els meus col·legues* ‘all (of) my colleagues’. *Tothom* and *tot el món* (or *tot lo món*) act as pronouns: *Tothom ho sap* ‘Everybody knows it’.

3.2.4 Partitive constructions

Partitive constructions are NPs denoting part of a whole or a group: *algun dels assistents* ‘some of the attendants’, *cinc dels meus llibres* ‘five of my books’, *moltes d’aquestes coses* ‘many of these things’, *cap de nosaltres* ‘none of us’. The quantifier

(*algun, cinc, molts, cap*) is the head of the construction and denotes a subset of the set denoted by the NP introduced by the preposition *de* ‘of’ (*els assistents, els meus llibres, aquestes coses, or nosaltres*). When the head is an inflected quantifier, it must agree in gender with the noun in the PP: *dues de les dones* ‘two of the women’, *dos dels homes* ‘two of the men’.

Some nouns with a quantitative meaning (*minoria* ‘minority’, *majoria* ‘majority’, *part* ‘part’, *subconjunt* ‘subset’, etc.) can appear as the head of partitive constructions: *una minoria de la gent* ‘a minority of people’, *la majoria del públic* ‘the majority of the public’. Partitive constructions and pseudopartitive constructions (described in 3.2.2: *dos litres de llet* ‘two liters of milk’) diverge in the sense that partitive constructions express a part-whole relationship.

When a partitive construction containing the collective nouns *majoria* ‘majority’ or *minoria* ‘minority’ and a PP with a plural NP occurs in the subject position, the verb can agree with either the head NP or the NP in the PP: *La majoria dels teus alumnes juga/juguen a futbol* ‘The majority of your students plays/play football’. And this is also the case with a noun of fraction: *La meitat de les dones aplaudeix/aplaudeixen* ‘lit. Half of the women claps/clap’.

4 Personal pronouns

Catalan personal pronouns belong to two classes: strong pronouns and weak, unstressed or clitic object pronouns (see 75.1 Word Classes, Inflectional Categories and Paradigms). First-person and second-person pronouns have a deictic nature. Generally, third-person pronouns are anaphoric and establish a semantic relation with an antecedent, as in *En Joan diu que ell no vindrà* (where *ell* = *en Joan*) ‘John says that he won’t come’. The form and function of strong pronouns are summarized in Table 3.

Table 3: Strong pronouns.

PERSON	FUNCTION	SINGULAR		PLURAL	
		MASCULINE	FEMININE	MASCULINE	FEMININE
1 st	Subject	jo		nosaltres	
	Prep. object	mi			
2 nd	Subject	tu		vosaltres	
	Prep. object				
3 rd	Subject	ell	ella	ells	elles
	Prep. object				
	Prep. object			si	

As shown in Table 3, strong pronouns may function as subjects or prepositional objects (15a,b). As prepositional objects, all of them may be reflexive, usually modified by the adjective *mateix* ‘-self’ (15c). Only *si* is exclusively reflexive (15d).

- (15) a. *Jo penso en tu.*
 ‘I am thinking of *you*’
 b. *Vosaltres desconfieu de nosaltres.*
 ‘*You*.PL are suspicious of *us*’
 c. *Sempre parles de tu (mateix).*
 ‘*You*.SG always talk about *yourself*’
 d. *La Carme sempre parla de si mateixa.*
 ‘Carme always talks about *herself*’

As mentioned in 1, in Catalan the subject can be dropped. However, strong pronouns are used to avoid ambiguity or to express emphasis (Rigau 1988). The presence of the pronoun *ella* avoids confusion in *He vist en Pere i la Maria, però ella no m’ha dit res* ‘I met Peter and Mary, but she didn’t talk to me’. In *Jo no hi estic d’acord* ‘I disagree with that’ the presence of the pronoun *jo* adds some emphasis that is not present in *No hi estic d’acord* ‘Disagree.1.SG with that’.

4.1 Weak personal pronouns

Weak or clitic personal pronouns are definite and mark either the accusative or dative case as shown in Table 4. For other variants and for a more detailed analysis of the form of weak pronouns see 5.1 Word Classes, Inflectional Categories and Paradigms.

Table 4: Weak pronouns.

PERSON	CASE	SINGULAR		PLURAL	
		MASCULINE	FEMININE	MASCULINE	FEMININE
1 st (reflexive and non-reflexive)	Accusative	em (me, m', 'm)		ens ('ns)	
	Dative				
2 nd (reflexive and non-reflexive)	Accusative	et (te, t', 't)		us	
	Dative				
3 rd (non-reflexive)	Accusative	el (lo, l', 'l)	la (l')	els ('ls)	les
	Dative	li		els ('ls)	
3 rd (reflexive)	Accusative	es (se, s', 's)			
	Dative				

Accusative pronouns act as direct objects. When the direct object pronoun refers to human entities, it can be reinforced by a strong form preceded by the preposition *a* (16a,b). Dative pronouns act as indirect objects and can also be reinforced (16c,d).

- (16) a. La Maria *em* convida (*a mi*).
 ‘Mary invites *me*’
 b. La Maria *ens* convida (*a nosaltres*).
 ‘Mary invites *us*’
 c. La Maria *em* regala un llibre (*a mi*).
 ‘Mary gives *me* a book’
 d. La Maria *ens* regala un llibre (*a nosaltres*).
 ‘Mary gives *us* a book’

The strong pronoun in (16) adds intensive emphasis. Dative pronouns are also used to express a part-whole relationship: *La Maria li renta les mans* ‘Mary washes her hands’. The dative pronoun *li* expresses the possessor and *les mans* expresses the part possessed. This dative pronoun is known as the possessive dative (see GIEC 2016, 682–684).

Weak reflexive pronouns can be reinforced by a strong pronoun modified by the adjective *mateix*. Plural pronouns can be used as reciprocal pronouns reinforced by the adverb *mútuament* ‘mutually’ or the expression *l’un a l’altre* ‘each other’ (see Todolí 2002, 1350–1356; GIEC 2016, 686–691).

- (17) a. Jo *em* rento (*a mi mateix*).
 ‘I wash *myself*’
 b. Ella *es* renta.
 ‘She washes *herself*’
 c. Elles *es* renten *l’una a l’altra*.
 ‘They wash *each other*’

4.2 Other weak pronouns

Catalan has three weak anaphoric pronouns that have neither the property of a person nor a definite character: *en*, *hi* and *ho*. However, they exhibit morphophonological behavior very similar to weak personal pronouns.

4.2.1 The pronoun *en*

The main syntactic functions of the weak pronoun *en* (or *ne*, *’n*, *n’*) are the following:

a) The pronoun *en* can refer to a bare NP acting as either the postverbal subject of an unaccusative predicate (8.3) or the direct object of a transitive verb: *Venen cotxes/En venen* ‘Some cars are coming/Some of them are coming’; *Jo menjo pomes/Jo en menjo* ‘I am eating some apples/I am eating some of them’. In the same way, the pronoun *en* can also represent a noun heading a NP with an indefinite specifier acting as either the postverbal subject of an unaccusative predicate or the direct object: *Venen molts cotxes/En venen molts* ‘Many cars are coming/Many of them are coming’; *Jo menjo tres pomes/Jo en menjo tres* ‘I am eating three apples/I am eating three of them’. When the indefinite NP contains a modifier (an AdjP or a PP), it can be represented by *en* or not. Given the sentence *Jo menjaré tres pomes vermelles* ‘I will eat three red apples’ we obtain *Jo en menjaré tres* (*en* = *pomes vermelles*) and *Jo en menjaré tres de vermelles* (*en* = *pomes*). Similarly, from *Jo menjaré una poma del cistell* ‘I will eat an apple from the basket’ we obtain *Jo en menjaré una* and *Jo en menjaré una del cistell*. On the use of the genitive preposition *de* in *Jo en menjaré tres de vermelles*, see 6.1e.

b) A PP headed by the preposition *de* ‘of’ can be represented by *en* when the PP acts as a complement of a noun in the following syntactic positions: direct object (18a), postverbal subject of an unaccusative verb (18b), and complement in a copular sentence (18c).

- (18) a. Coneixem tots els detalls *de la seva vida* > *En* coneixem tots els detalls.
 ‘We know all the details *of her life*’
 b. Només queda la pela *del plàtan* > Només *en* queda la pela.
 ‘Only the peel *of the banana* remains’
 c. Ella és la directora *de la coral* > Ella *n’és* la directora.
 ‘She is the director *of the choir*’

The pronoun *en* cannot be used when the noun denotes a part of an animate entity, as in *Rento la cara de la nena* ‘I wash the face of the girl’. The correct pronoun here is the dative pronoun *li*: *Li rento la cara* (**En rento la cara*). For the restrictions governing the presence of *en* as a nominal complement, see GIEC (2016, 700–703).

c) The pronoun *en* can represent a PP headed by the preposition *de* licensed by a verb: *parlar de* ‘talk/speak about’, *riure’s de* ‘to laugh at’, *oblidar-se de* ‘to forget’, *adonar-se de* ‘to realize’ *queixar-se de* ‘to complain of/about’, *recordar-se de* ‘to remember’, *sortir de* ‘to go out’, *tornar de* ‘to come/go back from’, etc.

- (19) a. En Pere parlava de política, però jo no *en* parlaré. (*en* = *de política*)
 ‘Peter spoke about politics, but I won’t speak about it’
 b. Abans no et queixaves de res i ara, de tot, te’n queixes. (*n* = *de tot*)
 ‘In the past you didn’t complain about anything and now you complain about everything’
 c. Jo vaig entrar al teatre quan tothom *en* sortia. (*en* = *del teatre*)
 ‘I went into the theater when everybody went out’

d) In some dialects, the complement of the copular verb *estar* can be represented by *en*: *En Pere està content, però en Joan no n'està* (or *n'hi està*) 'Peter is glad, but John isn't'.

4.2.2 The pronoun *hi*

The weak pronoun *hi* can represent any PP selected by the verb and headed by a preposition other than *de*, usually *a* 'to, in', *en* 'in', *amb* 'with' and *per* 'by' (*confiar en* 'to rely on', *insistir en* 'to insist on', *acostumar-se a* 'to get used', *adir-se amb* 'to match up with', *concordar amb* 'to agree with', *optar per*, 'to opt for', etc.):

- (20) a. Jo insistia *en la necessitat de fer-ho* > Jo *hi* insistia.
 'I insisted on the need to do it', 'I insisted on it'
 b. La corbata no s'adiu *amb la camisa* > La corbata no s'*hi* adiu.
 'The tie doesn't go with the shirt', 'The tie doesn't go with it'
 c. M'acostumaré *al cafè* > M'*hi* acostumaré.
 'I'll get used to having coffee', 'I'll get used to it'

The pronoun *hi* can also represent a directional PP selected by a verb of movement: *anar a* 'to go to', *arribar a* 'to arrive at/in', *entrar a/en* 'to go into', *pujar a* 'to go up', *passar per* 'to cross', etc: *Nosaltres anirem a Tarragona avui i ells hi aniran demà* 'We'll go to Tarragona today and they will go tomorrow'; *Alguns sortien del bar quan en Pere hi entrava* 'Some people left the bar when Peter went in'. The pronoun *hi* can also represent a stative locative PP or AdvP when the verb is copular or quasi-copular (21a–c) and a predicative complement (21d):

- (21) a. Soc *a casa* > *Hi* soc.
 'I am at home'
 b. Ja no s'estan *a Dénia* > Ja no s'*hi* estan.
 'They aren't in Dénia'
 c. Queda't *aquí* > Queda-t'*hi*.
 'Stay here'
 d. Arribaràs *ben cansada* > *Hi* arribaràs.
 'You will arrive vey tired'

In some dialects, *hi* can represent an AdjP acting as the complement in a copular sentence: *El nen està malat i el seu pare també hi està* 'The boy is ill and his father is too'. Moreover, in some dialects and in combination with other weak pronouns, *hi* can represent the indirect object instead of *li*: *Jo li donaré el llibre* 'I will give him/her the book' > *Jo l'hi donaré* 'I will give it to him/her'.

On the other hand, some adjuncts of the predicate can be represented by *hi*, for instance a PP/AdvP of manner and a locative adjunct: *No hi parla mai, amb educació/*

educadament ‘(S)he doesn’t ever speak politely’; *Hi vaig estudiar filologia, a la Universitat de Lleida* ‘I studied philology at the Universitat de Lleida’.

4.2.3 The pronoun *ho*

The weak pronoun *ho* represents a direct object formed by a demonstrative pronoun (*açò* ‘this’, *això* ‘this/that’ o *allò* ‘that’) or by a noun (or content) subordinate clause: *Si vols allò, t’ho regalaré* ‘If you want that, I will give it to you’; *Ahir va dir que no volia estudiar, però avui no ho ha dit* ‘Yesterday he said that he didn’t intend to study, but he didn’t say that today’; *Han fet tard, però no ho lamenten* (*ho* = *haver fet tard* ‘to have been late’) ‘They were late but they do not regret it’. Moreover, *ho* is the pronominal substitute par excellence of the complement in copular sentences: *Tu ets ric, però jo no ho soc* ‘You are rich, but I’m not (rich)’; *Estàs malalt o no ho estàs?* ‘Are you ill or aren’t you (ill)?’; *Em pensava que era tard, però no ho és pas* ‘I thought it was late, but it isn’t’; *Estàs de guàrdia o no ho estàs?* ‘Are you on call or not?’.

5 Adjectives and adjective phrases

Adjectives are lexical items that attribute a property to an entity or express its belonging to a class of entity. In general, they agree in gender and number with the noun that denotes the entity as *alt* in: *noi alt, noia alta, nois alts, noies altes* (lit. ‘boy tall, girl tall, boys tall, girls tall’). However, some adjectives are invariable: *fidel* is invariable in gender (*noia fidel* ‘loyal girl, *noi fidel* ‘loyal boy’), while *isòsceles* is invariable both in number and gender (*triangle/triangles isòsceles* ‘isosceles triangle/triangles’). Adjectives are the head of the AdjP and can appear with a specifier (an adverb of degree) and a PP as a complement: (*molt*) *fidel (a l’empresa)* ‘(very) loyal (to the company)’. Adjective phrases can occur within a NP, as modifiers of the noun (see Dols/Mansell 2017, 31–32): *una noia molt fidel a l’empresa* ‘a girl very loyal to the company’. Moreover, they can also be the complement in a copular sentence (*El meu veí és molt fidel a l’empresa* ‘My neighbor is very loyal to the company’) or a predicative complement (as *ben freda* in *Servirem la sopa ben freda* or *Servirem ben freda la sopa* ‘We’ll serve the soup very cold’).

Within the NP, the adjective usually appears in postnominal position (*una casa blanca* ‘a white house’). However, the adjective is prenominal when it has a non-restrictive meaning, as in *l’amable mestre* ‘the kind teacher’, where the adjective *amable* ‘kind’, in contrast with *el mestre amable*, does not restrict the reference of the NP. Adjectives such as *gran* ‘big’ and *pobre* ‘poor’, change their meaning according to their position in the NP, i.e. contrast *un gran home* ‘a great man’ and *un pobre home* ‘a miserable man’ with *un home gran* ‘an elderly man’ and *un home pobre* ‘a poor man’. Various classes of adjectives can be established:

a) Descriptive adjectives, which describe a quality of the noun: *content* ‘glad’, *petit* ‘small’, *tranquil* ‘calm’, *jove* ‘young’, *etern* ‘eternal’, etc. This quality can be inherent (or individual-level adjectives) or transitory (or stage-level adjectives) (Picallo 2002, 1646–1651; GIEC 2016, 548–551). Inherent adjectives may act as a complement in copular sentences with the verb *ser*, whereas transitory adjectives appear with the copular verb *estar*: *El pis és petit* ‘The apartment is small’, *El noi està content* ‘The boy is happy’. Some descriptive adjectives are, however, ambivalent (*avorrit* ‘boring’, *tranquil* ‘still/calm’, *nerviós* ‘nervous’, *inquiet* ‘restless’, etc.: *Ella és/està inquieta*). Most descriptive adjectives may occur with quantifiers: (*espai*) *bastant tranquil* ‘enough still (space)’, (*ciutat*) *massa petita* ‘too small (town)’. However, *etern* ‘eternal’ does not accept a quantifier.

b) Relational adjectives, which are associated to a noun (*alfabètic* < *alfabet*), classify the noun: *orgànic* ‘organic’, *alfabètic* ‘alphabetic’, *mensual* ‘monthly’, *fluvial* ‘fluvial’, *governamental* ‘governmental’, *francès* ‘French’, *papal* ‘papal’, etc. They do not express a quality but a relation between entities through the noun they modify and the noun which they stem from. They cannot occur in the prenominal position and cannot appear with quantifiers.

c) Other adjectives have a modal meaning (*possible* ‘possible’, *probable* ‘probable’, *segur* ‘sure’, etc.) or express a locative (temporal or spatial) meaning (*proper* ‘close’, *llunyà* ‘far’, *següent* ‘following’, *precedent* ‘precedent, previous’, etc.), or an aspectual meaning (*frequent* ‘frequent’, *continu* ‘continuous’, *ocasional* ‘occasional, temporary’, etc.). Focusing adjectives (*propi* ‘himself, etc.’, *simple* ‘simple’, *mer* ‘mere’, *únic* ‘only’, etc.) usually appear before the noun: *un simple jardiner* ‘a simple gardener’, *una mera coincidència*, ‘a mere coincidence’, etc.

d) Ordinal numerals (*primer* ‘first’, *segon* ‘second’, *tercer* ‘third’, *quart* ‘fourth’, *cinquè* or *quint* ‘fifth’, *dinovè* ‘nineteenth’, *vintè* ‘twentieth’, *trentè* ‘thirtieth’, etc.) are adjectives based on cardinal numbers (3.2.1a), and express priority or order among entities: *la primera casa* ‘the first house’, *la trentena setmana* ‘the thirtieth week’. It is very common to use cardinal numbers instead of ordinals above tenth: *fila 10* (or *fila deu*) ‘row 10’, *segle xx* (or *segle vint*) lit. ‘the twenty century’.

6 Prepositions and preposition phrases

Prepositions are invariable words that select a complement with which they make up a preposition phrase (PP), for instance the PP *sense tu* (‘without you’), consisting of the preposition *sense* (‘without’) and the pronoun *tu* (‘you’) as its complement. Prepositions differ clearly from adverbs, which are also invariable words but generally used without a complement (*tranquil·lament* ‘calmly’, *bé* ‘well’, *aquí* ‘here’). In Catalan, however, the distinction is not always clear: there are elements, such as the locative *dins* ‘inside’, which may select a complement, like prepositions (*Deixa les coses dins l’armari* ‘Put the things inside the wardrobe’), but they may also be used

with no complements, like adverbs (*Deixa les coses dins* ‘Put the things inside’). According to the most usual point of view in recent Catalan studies, these words will be analyzed as prepositions that may be used intransitively (Sancho Cremades 2002, 1697–1700; GIEC 2016, 719–720).

The complement of a preposition is typically a NP (*amb el teu germà* ‘with your brother’), but it may also be a finite clause (*sense que ho sàpiga ningú* ‘without anyone knowing about it’) or an infinitive clause (*sense dir-ho a ningú* ‘without telling anything to anybody’) and more unusually other elements, such as an adjective (*per bo* lit. ‘for good’ in *donar-ho per bo* ‘to consider it suitable’), an adverb (*fins ara* ‘until now’) or another PP (*per entre les pedres* lit. ‘by among the stones’). If the complement of the preposition is the first person singular pronoun, it is generally used in the oblique case (*amb mi* ‘with me’). But there are some exceptions, for instance, when the preposition *entre* (‘between, among’) selects a coordinate phrase including this pronoun: *entre tu i jo* ‘between you and me (lit. I)’, but *entre mi* ‘within myself’; or with certain prepositions with which the pronoun does not usually occur (*segons* ‘according to’, *malgrat* ‘in spite of’, or *mitjançant* ‘by means of’).

Catalan exhibits, on one hand, unstressed and stressed prepositions, and, on the other hand, simple and compound prepositions, besides a large number of prepositional locutions.

6.1 Unstressed prepositions

The simple prepositions *a* ‘to, in, on, at’, *en* ‘in, on, at’, *de* ‘of, from’, *amb* ‘with’ and *per* ‘for, by’, and the compound *per a* ‘for, to’ are unstressed. They are highly grammaticalized prepositions, with diverse functions (case-marking, prepositional regime supplement, location) and with a high frequency of use (Sancho Cremades 2002, 1730–1768; GIEC 2016, 725–751). The prepositions *a*, *de*, *per* i *per a* are contracted with the masculine singular and plural definite article: *al(s)*, *del(s)*, *pel(s)*, *per al(s)*, except for when the article is bound to the following vowel-initial word: *al pare* ‘to the father’ but *a l’avi* ‘to the grandfather’. The preposition *de* is always bound to the following vowel-initial and is written *d’* (*d’Andorra* ‘from Andorra’, *d’elles* ‘of/from them’).

a) The prepositions *a* and *en* may denote location and direction. The meaning of direction is generally expressed by *a*: *Anirem a Girona*, *a la plaça*, *a molts llocs* (‘We’ll go to Girona, to the square, to many places’). But the use of *en* is also possible before *un* ‘a’, *algun* ‘some’ and the demonstratives to avoid the contact of two vowels: *Anirem en un/algun/aquell hotel* (‘We’ll go to a/that hotel’). The expression of location exhibits more variation. In most frequent uses, *a* occurs before place name (*Viu a Girona* ‘(S)he lives in Girona’), before bare nouns with definite meaning (*Som a casa* ‘We’re at home’), before NPs with the definite article (*L’he vist a la plaça* ‘I’ve seen him in the square’) or with the interrogative *quin* ‘what/which’ (*A quina ciutat es troba?* ‘In

what city is (s)he now?'). Conversely, *en* appears with a bare NP with indefinite meaning (*Treballa en llocs diferents* '(S)he works in different places'), before a quantified NP (*La festa se celebrarà en qualsevol restaurant* 'The party will be celebrated in any restaurant') and with relative pronouns (*la ciutat en {la qual/què} es troba* 'The city in which (s)he is now').

b) Both prepositions are also used to denote temporal location. The preposition *a* is the most usual, especially when followed by the definite article (*Vam sopar a les 8* 'We had dinner at 8'; *Ens veurem a l'agost* 'We'll meet in August'), whereas *en* is more frequent in other contexts (*Ens veurem en uns dies o en Setmana Santa* 'We'll meet in a few days or at Easter') and may also have durative meaning (*No he parlat amb ell en els darrers mesos* 'I've not talked to him over the last few months').

c) The preposition *a* is used as a dative marker for the indirect object or for other dative objects: *Hem enviat una carta de protesta al director* 'We've sent a letter of protest to the director'; *Li han tallat els cabells a la Gal·la* 'Gal·la has her hair cut' (lit. 'They have cut her hair to Gal·la').

d) Although the direct object is not usually preceded by any preposition, there are some contexts where the use of the preposition *a* is either possible or obligatory. Generally, this is the case if the direct object is highly animate (or as animate as the subject) and/or in which the constituents are not located in their canonical position. The use of the preposition is compulsory with clitic doubling (22a), with the relative pronoun *qui* 'who/whom' (22b) and when the subject occurs after the verb or the verb is dropped (22c).

- (22) a. *Ens ha convidat a {nosaltres/tots els estudiants valencians}.*
 'They have invited {us/all the Valencian students}'
 b. *Convida a qui vulguis.*
 'Invite whoever you want'
 c. *Vam veure com perseguien els policies als lladres, i després el cotxe de policia al dels lladres.*
 'We saw the policemen chasing the thieves, and then the police car (chasing) that of the thieves'

Conversely, it is optional with any pronoun referring to people (23a) and when the direct object does not occur in its canonical position, as in dislocations (23b) or in focalizations (23c).

- (23) a. *Convidaran (a) tothom; No convidaran (a) ningú.*
 'They will invite everybody'; 'They won't invite anybody'
 b. *Encara no l'he vist avui, al/el meu pare.*
 'I've still not seen him today, my father'
 c. *AL/EL MEU PARE, veuré.*
 'MY FATHER, I will see'

e) The preposition *de* ‘of’ is used as a genitive marker (*el cotxe de Sergi* ‘Sergi’s car’, *casa de fusta* ‘house of wood’) and an ablative marker (*Han arribat de Barcelona* ‘They have arrived from Barcelona’). It is also widely used as a partitive marker; for instance, in partitive constructions (24a), in indefinite NPs, introducing the adjectival complement of a dropped noun (24b) or of a noun pronominalized by the partitive pronoun *en* (24c); introducing a dislocated constituent, especially if it is pronominalized by *en* (24d).

- (24) a. moltes *de* les meves amigues
 ‘many of my friends’
 b. No tinc cap bolígraf blau, però sí un *de* verd.
 ‘I don’t have any blue ball pens, but I do have a green one’
 c. –Quantes pomes tens? –En tinc tres *de* verdes.
 ‘–How many apples do you have? –I have three unripe ones’
 d. *De cafè*, *en* tens prou?
 ‘Coffee, do you have enough of it?’

f) The preposition *amb* expresses a relation of coincidence between two entities, as comitative (*Treballa molt a gust amb vosaltres* ‘I work at ease with you’), instrumental (*No tallis el pa amb el ganivet de la carn* ‘Don’t slice the bread with the meat cutting knife’) or manner (*Agafa les coses amb cura!* ‘Pick the things up with care’). There are in Catalan constructions such as *Amb la Roser anirem d’acampada a Aiguestortes* (‘With Roser, we’ll go camping to Aiguestortes’), where the comitative adjunct is left-dislocated and the plural covert subject includes the entity designated by the adjunct (Rigau 2002, 2062–2063; GIEC 2016, 737–738).

g) In formal language, the uses of *per* and *per a* are distinguished, but in most spontaneous registers only *per* is used in most of the linguistic territory, except for Valencian and some northwestern dialects (Solà 1987, 119–228; Sancho Cremades 2002, 1749–1755; GIEC 2017, 747–751, 1124–1126). The clearest difference between both prepositions is related to the causal meaning of *per* and the destination meaning of *per a*: *Tot ho fa per tu i per a tu* ‘(S)he does everything for you’ (*per*: ‘because of you’ and *per a*: ‘for your benefit’) (7.3 The Complex Sentence). The preposition *per* may express other meanings as ‘motion path’ (*Ves per l’autopista* ‘Drive by the motorway’), or the agent of periphrastic passives (*La llei ha estat aprovada pel Congrés* ‘The law has been passed by the Congress’). The preposition *per a* may have other meanings, such as temporal limit (*He d’acabar el treball per a demà* ‘I have to finish the work by tomorrow’), although the use of *per* is also common in most formal registers.

h) Most unstressed prepositions introduce prepositional regime supplements or adjuncts semantically close to the predicate, mainly *a* (*accedir a la vostra petició* ‘to accede to your request’), *en* (*pensar en vosaltres* ‘to think of/about you’) and *de* (*parlar de política* ‘to talk about politics’), and, to a lesser extent, *per* (*optar per una de les*

possibilitats ‘to opt for one of the possibilities’), *per a* (*servir per a la cuina* ‘to be useful/fit for the kitchen’) and *amb* (*Es conforma amb poca cosa* ‘(s)he settles for little’).

i) In the case of many verbs that select the preposition *en* when the regime supplement contains a NP, the preposition changes into *a* or *de* before an infinitive clause: *S’esforça en les coses que li manen* ‘(s)he puts effort into the things (s)he is required’, but *S’esforça a/de fer les coses que li manen* ‘(s)he strives to do what (s)he is required’. In formal language, this change is generalized to every use of *en* and *amb*: *Es conforma amb poc* ‘(S)he settles for little’, but *Es conforma a/de viure amb poca cosa* ‘(S)he settles for living with little’. For the elision of unstressed prepositions before the conjunctions *que* ‘that’, see 75.3 The Complex Sentence.

6.2 Stressed prepositions and prepositional locutions

Stressed prepositions may be simple (*sense* ‘without’, *entre* ‘between, among’, *contra* ‘against’, *durant* ‘during, for’, etc.) or compound (*cap a* ‘towards’, *fins a* ‘until, till, up to’, *com a* ‘as’ and *des de* ‘from, since’). Some of them may be used intransitively (*dins* ‘inside’, *davant* ‘in front, before, ahead’, etc.), as may prepositional locutions as well. In general, stressed prepositions exhibit more precise lexical meanings than unstressed ones, as in the corresponding English prepositions. The most notable uses of stressed prepositions are summarized below.

a) The preposition *sense* ‘without’ may be used without an explicit complement when it may be inferred from the same sentence or from a previous partial question: *–Vols el cafè amb sucre o sense?* –*Sense* ‘–Would you like the coffee with sugar or without? –Without’; *Passa’m uns quants fulls, que m’he quedat sense* ‘Pass me more paper sheets, since I don’t have any’.

b) The unstressed preposition *a* ending on some compound prepositions is sometimes dropped. This is the case for the preposition *cap a* ‘towards’, which denotes ‘direction’, when it occurs before a complement that may express this meaning on its own: *Cap a València* ‘towards València’, but *Vine cap aquí* ‘Come here’, *Ves cap amunt* ‘Go up’. In the preposition *fins a* ‘until, till, up to’, *a* is dropped mainly before a complement that may also be used with a spatial or temporal meaning without the unstressed preposition: *Fins al jardí* ‘up to the garden’, *fins a les festes de Nadal* ‘until Christmas’, but *fins allà* ‘up to there’, *fins fa poc* ‘until quite recently’, *fins ara* ‘until now’.

c) Certain space prepositions may be used intransitively if the complement is inferable from the linguistic or the situational context: *Obre l’armari i deixa la roba dins (l’armari)* ‘Open the wardrobe and put the clothes inside (the wardrobe)’. In many occasions, these prepositions establish binary oppositions related to different spatial axes: *davant* ‘before’/*darrere* ‘behind’ (horizontal axis); *damunt* ‘above, on, over’/*davall* ‘below, beneath, underneath’ (vertical axis); *dins* ‘inside’/*fora* ‘outside’ (inside-

ness); (*a*) *prop* ‘near, close to’/*lluny* ‘far away’ (proximity), etc. If the complement is explicit, the unstressed linking preposition *de* ‘of’ is necessary in the case of *fora*, *prop* and *lluny*, but optional with the rest: *Deixa-ho davant (de) la casa* ‘Lay it in front of the house’, but *lluny de la casa* ‘far away from the house’. If the complement is a personal pronoun, it may either be introduced by *de* (*davant de mi* ‘in front of me’) or adopt the possessive form (*davant meu*, lit. ‘in front mine’). It may also be a dative pronoun in possessive dative constructions, as *li* in *Se li va posar davant* ‘(S)he stood before him/her’. The spatial meaning of these prepositions may be reinforced by means of *a* (*a dalt de l’armari* ‘on the top of the wardrobe’) and sometimes with *al* (*al davant* ‘at the front’, *al darrere* ‘at the back’, *al damunt* ‘on/at the top’, *al davall* ‘at the bottom’).

d) Preceded by *a*, the prepositions *sobre* ‘on, over’ and *sota* ‘under, below’ are equivalent to *damunt* and *davall*, respectively, and then they can also be used intransitively: *Deixa-ho a sota ((de) l’armari)* ‘Put it below (the wardrobe)’, but only *Deixa-ho sota l’armari* ‘Put it under the wardrobe’.

e) The prepositions *abans* ‘before’ and *després* ‘after, afterwards’, which denote temporal location anterior and posterior to the reference point respectively, may also be used intransitively: *Segur que arriben abans/després* ‘They will arrive before/afterwards for sure’. They may be followed by a NP or an infinitive clause introduced by *de*, or a finite clause with *que*: *després {de la festa/de sopar/que sopàrem}* ‘after {the party/having dinner/we had dinner}’.

f) The exceptive particles, usually analyzed as prepositions, are derived from participles (*excepte*, *tret*, *llevat* ‘except for’), quantifiers (*menys* ‘less’) or locatives (*fora*, lit. ‘out’). They indicate that there are elements excluded from a group: *Vindran tots {excepte/tret de} la teva companya* ‘Everybody will come except for your colleague’ (GIEC 2016, 768–769).

g) Most prepositional locutions occur in the pattern (preposition) + noun + preposition: *a condició de* ‘with the condition of’, *a causa de* ‘because of’, *en relació amb* ‘in relation to’, *gràcies a* ‘thanks to’, etc. The noun sometimes occurs with the definite article: *al costat de* ‘beside’, *a la vora de* ‘close/near/next to’. Instead of the noun, a non-finite verbal form may sometimes also occur: *tocant a* ‘with regard to’, *a partir de* ‘from, on the basis of’, (*en*) *acabat de* ‘just after’. As in the case of stressed prepositions, the complement may remain implicit in some locutions, and then they are used without the last preposition: *Deixa la maleta al costat (de la porta)* ‘Put the suitcase aside (beside the door)’.

7 Adverbs and adverb phrases

Adverbs are invariable words that may modify a verb (*bé* ‘well’ in *parla bé* ‘(s)he speaks well’), an adjective (*molt* ‘very’ in *molt bonic* ‘very nice’) or another adverb (*ben* ‘rather, quite, fairly’ in *ben tranquil·lament* ‘rather calmly’). They may also have an

effect on the whole sentence (such as *probablement* ‘probably’ in *Probablement arribaran demà* ‘Probably they will arrive tomorrow’) and less commonly on other syntactic constituents (such as *només* ‘only’, which modifies a PP in *només amb tu* ‘only with you’ and a NP in *només el peix* ‘only fish’).

Formally, the most productive class of adverbs is that of adverbs ending with *-ment* ‘-ly’, in which the constituent *-ment* is added to an adjectival lexeme in the feminine form: *tranquil·la* + *ment* ‘calm (f.) + ly’ > *tranquil·lament* ‘calmly’. As in compound words, the main stress is on the second constituent, but the adjectival lexeme keeps a secondary stress, which is spelled as it is spelled in the feminine form of the adjective: *pràctica* ‘practical’ (adjective f.) and *pràcticament* ‘practically’ (adverb). If several adverbs with *-ment* are coordinated, the final constituents may be suppressed in all the adverbs except for the first one: *Tot ho resol ràpidament i eficaç* ‘(S)he solves everything quickly and efficiently’ (lit. ‘quickly and efficient’). In modern-day Catalan, however, keeping the ending on all the conjuncts is the most common usage: *Tot ho resol ràpidament i eficaçment*.

So-called adjectival adverbs are also derived from adjectives, but without any deriving suffix. Thus, the form of the adverb coincides with that of the adjective in masculine. For instance, *alt* ‘tall’ is an adjective in *un nen alt* ‘a tall child’ but an adverb in *parlar alt* ‘to speak aloud’ (lit. ‘to speak tall’). Like in English, there are relatively few adverbs with *-ment* and adjectival ones that derive from the same adjective: *camina lent/lentament* ‘(S)he walks slow/slowly’, *s’expressa clar/clarament* ‘(S)he speaks clearly’. Most adjectival adverbs occur with a closed set of verbs or constitute more or less lexicalized constructions: *parlar clar* ‘to talk clearly’, *treballar fort* ‘to work hard’, *filar prim* ‘to quibble, to split hairs’, etc.

Besides these deadjectival adverbs, there are others related to less clearly defined subcategories. Specifically, these include demonstrative adverbs (*així* ‘so/thus’), relative and interrogative adverbs (*on* ‘where’, *quan* ‘when’, *com* ‘how’) or quantifiers (*molt* ‘very, much, many’, *poc* ‘a little, a few’, *massa* ‘too (much/many)’).

As in the case of prepositions, there are a lot of adverbial locutions. Many of them have the form of a PP (*a les palpentes* ‘blindly’, *en principi* ‘in principle’, *per ara* ‘by now’), sometimes with correlative prepositions (*de bat a bat* ‘widely’, *de gom a gom* ‘crammed full, full to the brim’). Those based on adverbs (*ben bé* ‘exactly’, *ara per ara* ‘by now’, *mai de la vida* ‘never’) or on participles (*tot seguit* ‘just now’, *comptat i debatut* ‘all in all’) are less frequent. There are also some rather fixed ones, fully lexicalized and written as a single word (*només* ‘only’, from *no* ‘not’ and *més* ‘more’; *potser* ‘maybe’, from *pot ser* ‘may be’; *tothora* ‘everytime’, from *tota hora* ‘every time’).

Generally speaking, adverbs are the only constituent of the adverb phrase (AdvP), though the phrase may also consist of specifiers and, sometimes, complements. The AdvP may contain another adverb or an equivalent expression as a specifier; for instance, *molt* ‘very’ in *molt bé* ‘very well’ or *molt lentament* ‘very slowly’. The adverbs derived from adjectives that select a complement acquire this syntactic property as

well. Thus, they may select a complement that, like those of the corresponding adjectives, is introduced by an unstressed preposition: *contràriament a la vostra opinió* (like *contrari a la vostra opinió* ‘contrary to your opinion’), *diferentment de tu* (like *diferent de tu* ‘different from you’).

As for the meaning and the type of constituent that they modify, they are classified as adverbs of manner (*tranquil·lament* ‘calmly’, *bé* ‘well’, *alt* ‘aloud’), place (*amunt* ‘above, upward’, *allà* ‘there’), time (*ara* ‘now’, *sempre* ‘always’, *aviat* ‘soon’) and aspect (*ja* ‘already’, *mensualment* ‘monthly’), degree (*molt* ‘very, much, many’, *poc* ‘a little, a few’, *gens* ‘any, no’), modality (*potser* ‘maybe, perhaps’, *necessàriament* ‘necessarily’), focusing adverbs (*només* ‘only’, *precisament* ‘precisely, exactly’) and polarity adverbs (*sí* ‘yes’, *no* ‘no’). The main features of each class are summarized below. As for polarity adverbs, see 10.

a) Adverbs of manner denote how the situation expressed by a predicate is performed. Most of them are adverbs with *-ment* or adjectival adverbs, but they also include the relative and interrogative *com* ‘how’, the demonstrative *així* ‘so, thus’ and the comparatives *millor* ‘better’ and *pitjor* ‘worse’, some formally heterogeneous adverbs (*bé/ben* ‘well’, *corrents* ‘fast, hastily’, *dempeus* ‘up, upright’, etc.) and a wide set of adverbial locutions (*a poc a poc* ‘little by little, step by step’, *a les palpentes* ‘blindly’, *de genollons* ‘on one’s knees’, etc.).

b) Adverbs of manner usually qualify the verbal predicate like an adjective qualifies a noun. Thus, it may be said that the activity of *caminar* ‘to walk’ (verb) is done *lentament* ‘slowly’ (adverb) in the same way that it is said that a *caminada* ‘walk’ (noun) is *lenta* ‘slow’ (adjective). For this reason, adverbs of manner are normally included within the VP (25a). However, they may also function as sentential adverbs when they qualify the subject (25b), one of the participants in the speech act (25c) or the whole sentence (25d).

- (25) a. El professor va reconduir *prudentment* i *honradament* el debat.
 ‘The teacher redirected the debate *cautiously and honestly*’
 b. El professor, *prudentment*, va reconduir el debat.
 ‘*Cautiously*, the teacher redirected the debate’
 c. *Honradament*, no en sabia res.
 ‘*Honestly*, I didn’t know anything about it’
 d. *Sortosament*, l’operació va ser un èxit.
 ‘*Fortunately*, the operation was a success’

Adverbs of manner may also modify an adjective: *tristament cèlebre* ‘sadly notorious’, *inexplicablement desconsiderat* ‘unexplainably ruthless’. Nevertheless, in such uses, they often add a quantitative meaning (‘very notorious/ruthless’).

c) However, not all adverbs of manner act semantically as qualifiers. There are also some that derive from adjectives of relation and introduce classifications (§ 5b): *analitzar sintàcticament una oració* ‘to analyze a sentence syntactically’ (by means of

una anàlisi sintàctica ‘a syntactic analysis’), *classificar alfabèticament uns noms* ‘to classify some nouns alphabetically’ (by means of *una classificació alfabètica* ‘an alphabetical classification’). Some of these adverbs may also act as sentential adverbs and denote the point of view from which the sentence should be interpreted or evaluated: *Sintàcticament, l’anàlisi és incorrecta* ‘From a syntactic point of view, the analysis is wrong’; *Políticament, és un error no acceptar la proposta* ‘Politically, it is a mistake not to accept the proposal’. Finally, we should highlight the existence of adverbs that require a plural context and express reciprocity (such as *mútuament* ‘mutually’ in *S’acaronen mútuament* ‘They caress mutually’) or impose the interpretation of a group action (such as *conjuntament* ‘together’ in *L’ajuntament i la universitat han organitzat conjuntament l’acte* ‘The council and the university have organized the event together’) or a distributive meaning (such as *respectivament* ‘respectively’ in *L’un i l’altre treballen, respectivament, en la redacció i la correcció dels textos* ‘One and the other work in writing and correcting texts, respectively’).

d) Among adverbs of place, there are demonstratives (*ací* ‘here’, *aquí* ‘here, there’), those denoting quantification (*arreu* ‘everywhere’, *per tot* ‘all over, throughout’, *enlloc* ‘anywhere, nowhere’) and those expressing relation (*avant* ‘ahead’, *arrere* ‘back’), as well as the relative and interrogative *on*. The demonstrative adverbs set up space deictic oppositions similar to those expressed by nominal demonstratives (see § 3.1b). As in that case, they exhibit a two-term system in most Catalan dialects: *aquí* ‘here’ (proximal)/*allà* o *allí* ‘there’ (distal); but a ternary system in Valencian and some northwestern dialects (Payrató 2002, 1165–1169; Nogué 2015, 207–210; Pérez Saldanya 2015). In the three-term system, the most used forms are: *ací* o *aquí* ‘here’ (immediate proximity or proximity to the addressor), the Castilian-influenced *ahí* ‘there’ (mediate proximity or proximity to the addressee) and *allí* o *allà* ‘there’ (distance from both the addressor and the addressee). In the expression of distance, *allí* denotes a more precise location and *allà* a fuzzier one.

Demonstrative adverbs may express location (*Viu aquí* ‘(S)he lives here’) or destination (*Ves-te’n allà* ‘Go there’), depending on the verb, and they may act as complements of dynamic space prepositions (*des d’ací* ‘from here’, *per allí* ‘over there’, *cap allà* ‘(to) there’). Like nominal demonstratives, they may be modified by *mateix* ‘itself’ (*aquí mateix* ‘just here’ lit. ‘here itself’).

e) The adverbs *endavant* (o *avant*) ‘ahead’, *endarrere* (o *arrere*) ‘back’, *amunt* ‘up’, *avall* ‘down’ express a dynamic space relation and denote direction and orientation. For this reason, they occur with verbs expressing these meanings and with the preposition *cap* (a) ‘to, towards’: *Aneu endavant* ‘Go ahead’; *Mireu cap amunt a la paret* ‘Look up on the wall’. These adverbs may appear with a degree quantifier (*molt avall* ‘far down’, *una mica endavant* ‘a bit further ahead’) and may be followed by a complement introduced by *de*, which depends on the quantifier: *més avall d’Alacant* ‘further down from Alacant’, *tan endins de la cova com l’altre dia* ‘as far inside the cave as the other day’. They may also be used after a bare noun or a NP in constructions with various meanings (Bartra Kaufmann/Suñer 1992; Pérez Saldanya/Rigau 2007;

GIEC 2016, 803–805). For example, denoting the path followed on a route (26a) or how a body part is oriented (26b).

- (26) a. Se'n va anar corrents {[carrer *amunt*]/[pel passeig de Gràcia *amunt*]}.
 '(S)he left hastily {[*up* the street]/*up* the passeig de Gràcia}]'
- b. Dormia {[panxa *avall*]/[de panxa *avall*]/[amb la panxa *avall*]}.
 '(S)he slept face *down*'

These adverbs have as their static correlates the locatives *davant* 'in front, before', *darrere* 'behind', *damunt* 'above', *davall* 'below', analyzed in 6.3c as intransitive prepositions. The adverb *baix* also has a static meaning, but it selects no complement: *Espera'm baix* 'Wait for me downstairs'.

f) Adverbs of time encompass or locate temporally the situation expressed by the predicate. They may occur in different positions within the sentence, as in the case of *ara* 'now' in the following examples: *Ara mon pare es troba més bé* 'Now my father feels better'; *Mon pare, ara, es troba més bé* 'My father, now, feels better'; *Mon pare es troba més bé ara* 'My father feels better now'. Moreover, a sentence may contain more than one temporal adverb or expression, but the order of occurrence must show the order of inclusion: *Demà arribarem a les 8* 'Tomorrow we will arrive at eight'; *Arribarem demà a les vuit* 'We will arrive tomorrow at eight'. Among adverbs of time, there are the relative and interrogative *quan*, the deictic adverbs (*ara* 'now', *demà* 'tomorrow', *demà passat* 'the day after tomorrow', *enguany* 'this year', *actualment* 'presently, nowadays', etc.), the anaphoric adverbs (*llavors* or *aleshores* 'then') and those denoting quantification (*sempre* 'always', *mai* 'never'), simultaneity (*mentrestant* 'meanwhile') or temporal distance in relation to a temporal point of reference (*tard* 'late', *aviat* 'early').

g) Aspectual adverbs refer to the internal time of the situation denoted by the predicate; that is, they express duration (*breument* 'briefly, shortly', *momentàniament* 'momentarily'), repetition or frequency (*diàriament* 'daily', *normalment* 'usually', *sovint* 'often') or the stage in the development of a situation (*ja* 'already', *encara* 'still, yet', *completament* 'completely', *del tot* 'fully, entirely'). The adverbs *ja* and *encara* usually occur before the verb (*Ja s'ha adormit* '(S)he is already sleeping'), whereas the other adverbs denoting stage appear after the verb (*Va revisar completament la traducció* '(S)he completely revised the translation'). The other aspectual adverbs are usually placed after the verb, but other positions in the sentence are also possible: *Passegen sovint per la platja* 'They often walk along the beach'; *Normalment arriben puntuals* 'Normally, they arrive on time'.

h) When variable, adverbs of degree coincide with quantifiers in the masculine singular form (7.5.1 Word Classes, Inflectional Categories and Paradigms): *molt* 'much, many', *poc* 'a little, a few', *gens* 'any, no', *massa* 'too much/many', *més* 'more', etc. They denote the value with which a certain feature or situation is presented on a degree scale. They may modify a verb (*Parla massa* '(S)he talks too

much'), an adjective (*poc dolç* 'a little sweet'), an adverb (*més lentament* 'more slowly') and some PPs (*molt per sota de les seues possibilitats* 'well below their possibilities'). The properties of the various types of degree adverbs coincide with those of quantitative words and are analyzed in 3.2.2.

i) Adverbs of modality express the possibility (*potser* 'maybe, perhaps', *segurament* 'surely', *tal vegada* 'perhaps, probably'), necessity (*necessàriament* 'necessarily', *per força* 'obligatory, compulsory') or degree of truth (*realment* 'really', *de debò* 'actually, truly', *segons sembla* 'apparently, seemingly') of an idea expressed in the sentence. Therefore, they are sentential adverbs. They may appear at the beginning of the sentence (*Segurament haurem de demanar ajuda*, 'Surely, we will have to ask for help'), but they may also occur at the end or somewhere in between (*Ens contestaran demà, sens dubte* 'They will answer us, without a doubt'; *Vindran possiblement demà* 'They will possibly come tomorrow'). Some adverbs denoting a high degree of truth may occur in emphatic constructions linked to the sentence that they modify by means of the conjunction *que*: *Naturalment que els vam avisar* 'Of course/Naturally, we did warn them'.

j) Focusing adverbs highlight the informative relevance of the constituent to which they are adjoined and specify it or contrast it in relation to other possible or presupposed constituents. They may indicate exclusion (*només, sols* 'only', *exclusivament* 'exclusively', *tot just* 'just'), positive inclusion (*fins i tot* 'even', *també* 'also, too, as well') or negative (*ni tan sols* 'not even', *tampoc* 'neither'), specificity (*en concret* 'specifically', *exactament* 'exactly', *precisament* 'precisely, just', *sobretot* 'mainly, mostly') and approximation (*almenys, si més no* 'at least/most', *gairebé* 'almost, nearly'). Due to their meaning, they may affect every sentential constituent, including subordinate clauses; for example, a NP (*fins i tot tu* 'even you'), an AdjP (*una relació exclusivament epistolar* 'an exclusively epistolary relationship'), a VP (*només plora* '(S)he only cries'), a PP (*sobretot per aquesta raó* 'mainly for this reason'), an AdvP (*precisament demà* 'just tomorrow'), a conditional clause (*només si ens ajudes* 'only if you help us'), etc.

8 Verbs and verb phrases

Sentences express a predicate in a temporal frame. Prototypically, the predicative meaning is displayed by a verb, which also expresses the temporal frame by means of its inflective properties, as in (27a), or through an auxiliary verb, as in (27b).

- (27) a. Normalment *plou* a l'agost.
 'It usually *rains* in August'
 b. Aquesta setmana *ha plogut* molt.
 'This week it *has rained* a lot'

Verbal tenses are studied in section 9 and verbal periphrases in section 10. Meanwhile, copular and quasi-copular verbs are described in 8.4.

Verbal predicates determine the arguments of the sentence. An argument is internal if it is a constituent of the VP together with the verb. An external argument is a constituent of the sentence. Usually, the external argument acts as the subject of the sentence, which may be silent, as shown in section 1. Some verbs express the predicate without any arguments, such as the meteorological verbs *ploure* ‘to rain’, *nevar* ‘to snow’, *tronar* ‘to thunder’, *pedregar*, ‘to hail’, etc. A syntactic way to avoid the external argument is the use of the clitic *es*: *Ara es viatja molt* ‘Nowadays people travel a lot’, *S’han venut força llibres* ‘A lot of books have been sold’ (where the internal argument agrees with the verb).

Three wide types of verbal predicates must be established: transitive, intransitive and unaccusative verbs or VPs. Transitive verbs select a direct object: *Mengen patates* ‘They eat potatoes’, where *potatoes* is an internal argument acting as a direct object. Intransitive verbs do not permit a direct object: *La Maria salta* ‘Mary is jumping’. Unaccusative verbs share some properties with transitive verbs and other properties with intransitive verbs: *Cauen fulles* ‘Some leaves are falling’.

8.1 Transitive verbs

Prototypically, transitive verbs select two arguments: an external argument – the sentential subject – and an internal argument – the direct object: *En Pere llegeix una novel·la* ‘Peter reads a novel’. The direct object of the verb *llegir* can be implicit: *En Pere llegeix* ‘Peter reads’. However, other transitive verbs such as *conèixer* ‘to know’, *recuperar* ‘to recover’, *lamentar* ‘to regret’ do not permit an implicit object (GIEC 2016, 833–835).

a) Some transitive verbs accept the passive construction with the verb *ser/ésser*, where the internal argument acts as the subject and the external argument is not present or it is expressed as a PP headed by the preposition *per*: *L’ermita fou reconstruïda (pels vilatans) el segle XVIII* ‘The chapel was restored (by the inhabitants) in the 18th century’, or with the clitic *se*: *Es va reconstruir l’ermita el segle XVIII* ‘The chapel was restored in the 18th century’ (see Bartra Kaufmann 2002, 2126–2135; Alsina 2016, 377–379).

b) There are some transitive verbs, such as *fer* ‘to do/make’ and *donar* ‘to give’, that are known as light verbs, because their meaning strongly depends on the internal argument, and they usually can be paraphrased using an intransitive verb: *fer salts/saltar* ‘to jump’, *fer badalls/badallar* ‘to yawn’, *fer broma/bromejar* ‘to joke’, or the meteorological impersonal predicates *fer vent/ventejar* ‘to blow’, *fer llampecs/llampejar* ‘to flash (lightning)’, etc. However, the verb *donar* and its first internal argument can be paraphrased using a transitive verb: *donar consol a algú/consolar algú* ‘to comfort’. On light verbs, see Rosselló (2002, 1880–1882) and GIEC (2016,

825–833). On measure verbs *costar* ‘to cost’, *pesar* ‘to weigh’, *durar* ‘to last’, etc., whose direct object denotes a measure of price, weight, time, etc. (*Això costa 20 €* ‘It costs 20 €’), see GIEC (2016, 836–837). On the presence of the preposition *a* heading a direct object, see 6.1d.

c) Some transitive verbs (*donar* ‘to give’, *regalar* ‘to give’, *enviar* ‘to send’, *dir* ‘to say’, *prometre* ‘to promise’, *manar* ‘to order’, *prendre* ‘take out’, etc.) select a second internal argument that expresses ‘goal’ and acts as indirect object headed by the preposition *a* (6.1c) or represented by a dative weak pronoun (4.1): *El professor va anunciar la notícia als estudiants* ‘The professor announced the news to his students’.

d) Transitive verbs such as *convèncer de* ‘to convince’, *convidar a* ‘to invite’ obligar *a*, *basar en*, etc. select as a second internal argument a PP that does not act as an indirect object but as a prepositional regime supplement. Thus, in *Vaig convèncer la Maria de la meva innocència* ‘I convinced Mary of my innocence’, the PP *de la meva innocència* can be represented by the pronoun *en* (*En vaig convèncer la Maria*), likewise the PP *a la festa* in *Van convidar la Maria a la festa* ‘They invited Mary to the party’ can be substituted by the pronoun *hi* (*Hi van convidar la Maria*). The second internal argument can have a locative meaning. In this case, the preposition is not fixed and the argument can be expressed by a locative adverb, as in *Posarem allò a l’armari/dins l’armari/allà* ‘We’ll put that in the cupboard/inside the cupboard/in there’.

8.2 Intransitive verbs

Intransitive (or unergative) verbs do not select a direct object. Some of them only select an external argument, such as the verbs of activity *saltar* ‘to jump’ *lladrar* ‘to bark’, *caminar* ‘to walk’, *dormir* ‘to sleep’, *plorar* ‘to cry’, *viure* ‘to live’, *sopar* ‘to have dinner’: *Els gossos lladren* ‘The dogs are barking’. Moreover, some verbs expressing the production of liquids, sounds or light are intransitive: *suar* ‘to sweat’, *sagnar* ‘to bleed’, *cruixir* ‘to rustle’, *brillar* ‘to shine’, *titil·lar* ‘to twinkle’ (and some impersonal verbs such as *tronar* ‘to thunder’ and *ploure* ‘to rain’): *Les fulles seques cruixien i les estrelles titil·laven* ‘The dry leaves rustled and the stars twinkled’.

Some intransitive verbs of action such as *telefonar* ‘to phone’ or *pegar* ‘to hit’ select an indirect object: *Li telefonaré demà* ‘I’ll phone him tomorrow’; *Peguen al gos* ‘They hit the dog’.

Perception verbs are transitive verbs. However, verbs of involuntary perception (*veure* ‘to see’, *sentir* ‘hear’, etc.) are intransitive when the pronoun *hi* is lexically incorporated: *veure-hi*, *sentir-hi*. For this reason, a sentence like *En Pere no hi sent* means ‘Peter is not able to hear, because he is (rather) deaf’.

There are some verbs that select a complement but it is not treated as a direct object. This complement is usually a PP acting as a prepositional regime supplement: *dependre de* ‘to depend on’, *confiar en* ‘to rely on’, *parlar de/sobre* ‘to speak about’,

etc. Frequently, the verb is pronominal: *adonar-se de* ‘to realize’, *equivocar-se de* ‘to be wrong about’, *tendir a* ‘to tend to’, *versar sobre* ‘to deal with’: *Ella no s’adona de l’error* ‘She doesn’t realize the mistake’. Moreover, some pronominal verbs correlate with a transitive non-pronominal verb: *aprofitar-se de/aprofitar* ‘to take advantage of’, *recordar-se de/recordar* ‘to remember’, *caracteritzar-se de/caracteritzar* ‘to characterize’, etc. (Rosselló 2002, 1942–1946).

Sometimes an intransitive verb appears with a cognate object, a sort of direct object semantically related to the meaning of the verb: *viure la vida plenament* ‘to live life to the fullest’, *ballar un ball/un vals* ‘to dance a dance/a waltz’.

8.3 Unaccusative verbs

There is a series of sentences, apparently intransitive, whose subject is semantically interpreted as an internal argument, as a theme or patient. Their verb does not select an external argument and is unable to treat its internal argument as a direct object. Consequently, the internal argument acts as the grammatical subject of the sentence. Precisely because of that, they are called “unaccusative (or ergative)” verbs. In *Ha nascut una criatura* ‘A child is born’, *una criatura* is not interpreted as the agent of an action, but as a patient.

a) Unaccusative verbs express: (i) change of state (*néixer* ‘to be born’, *morir* ‘to die’, *aparèixer* ‘to appear’, *desaparèixer* ‘to disappear’, etc.), (ii) that something is taking place (*ocórrer* ‘to occur’, *passar* ‘to happen’, *succeir* ‘to take place’, etc.), (iii) shortage or remnants (*faltar* ‘to lack’, *sobrar* ‘to be left over’, *quedar* ‘to remain’, etc.), and (iv) the existence of an obligation or necessity (*caldre* ‘to be required’, *tocar* ‘to be somebody’s turn’, *urgir* ‘to be urgent’, etc.), among others.

Their behavior shows that unaccusative verbs are hybrids of transitive and intransitive verbs. On the one hand, their subject tends to appear in postverbal position, the typical position of verbal objects: *Falten les nostres signatures* ‘Our signatures are needed’. On the other, the postnominal subject can be a bare NP, like direct objects as in *Sempre passen coses* ‘Some things always happen’. Conversely, the subject of a transitive or an intransitive verb cannot be a bare NP. Further, the subject of an unaccusative verb can be pronominalized with *en* under the same conditions as a direct object: *De diners, en calen molts* ‘As for money, one requires a lot’. In addition, the internal argument of an unaccusative or a transitive verb expressing an achievement can appear as an adjunct of an absolute participle clause: *Passada la guerra ...* ‘After the war...’.

b) Besides the unaccusative verbs, there are some syntactic constructions with unaccusative properties. As observed by Fabra (1956, 37–40, 85–87), verbs of movement (*venir* ‘to come’, *arribar* ‘to arrive’, *tornar* ‘to come back’, *entrar* ‘to go into’, etc.) are ambivalent. They can display unaccusative behavior, as in (28a) with a postverbal bare NP that can be pronominalized using *en*, or intransitive behavior, as in (28b) with

an agentive subject allowing an adjunct of purpose. On these constructions, see Gràcia i Solé (1989), Rigau (2013), Rosselló (2002, 1887–1895), Solà (1973, 15; 2009, 298–302).

- (28) a. Venen pluges/En venen.
 ‘Rains are coming’
 b. Nosaltres havíem vingut per visitar-la.
 ‘We had come in order to visit her’

c) Other locative constructions with the locative pronoun *hi* exhibit unaccusative properties although the verb is not unaccusative, such as (29).

- (29) En aquell taller *hi* treballaven aprenents.
 In that garage *locative clitic* worked.3.PL apprentices
 ‘Some apprentices worked in that garage’

Note that the verb *treballar* is an intransitive verb. However, in this construction it behaves as an unaccusative verb allowing a bare NP, which can be pronominalized by *en*: *En aquell taller n’hi treballaven*. These constructions have an existential flavor similar to sentences with the predicate *haver-hi* (e.g., *Hi ha aprenents que treballen en aquell taller* ‘There are some apprentices working in that garage’). On the predicate *haver-hi*, see 8.4h. On the properties of these constructions and others related to them, see GIEC (2016, 850–851), Rigau (1997), Rosselló (2002, 1894–1895), and Solà (1994).

d) Some verbs that express the production of a liquid (*vessar* ‘to pour’, *traspuar* ‘to exude’, *rajar* ‘to spring up’, etc.) can exhibit unaccusative behavior in constructions with a source locative (or a dative): *Ell s’adonà que vessava aigua del safareig* ‘He realized that some water was pouring from the sink’.

e) Causative verbs, which are transitive, can correlate with a pronominal verb with unaccusative properties: *trencar/trencar-se* ‘to break’, *enfonsar/enfonsar-se* ‘sink’, *purificar/purificar-se* ‘purify’, etc. The causative construction *El vent trencà el vidre* ‘The wind broke the pane’ correlates with the unaccusative constructions *El vidre es trencà*, *Es trencà el vidre* ‘The pane broke’, where the internal argument acts as the sentential subject. Some of these verbs are semantically classified as psychological: *avergonyir/avergonyir-se* ‘embarrass/to be embarrassed’, *emocionar/emocionar-se* ‘to thrill/to be thrilled’, *preocupar/preocupar-se* ‘to worry/to be worried’: *La notícia emocionarà la Maria/La Maria s’emocionarà* ‘The news will move Mary/Mary will be moved’. Other psychological verbs, which do not correlate with a causative verb, also exhibit unaccusative properties (i.e., the sentential subject is interpreted as the theme), such as *agradar* ‘to like’, *doldre* ‘to hurt’, *plaure* ‘enjoy’, etc.): *Al director no li agraden les crítiques* ‘The director doesn’t like criticisms’ (GIEC 2016, 853–855; Rosselló 2002, 1921–1927).

8.4 Copular and quasi-copular verbs

In copular and quasi-copular sentences, the element supporting the predication is not the verb but a non-verbal phrase – the complement –, such as the AdjP *molt bonic* in *Això és molt bonic* ‘This is very nice’. The copular verb *ser* ‘to be’, or the oldest variant *ésser*, only provides the sentence with information about tense, person and number, and it is the link between the subject and the predicate. However, quasi-copular verbs such as *estar* ‘to be’, *continuar* ‘to continue’, *esdevenir* ‘to become’, *quedar(-se)*, ‘to remain’, etc. add aspectual information, while the quasi-copulars *semblar* ‘to seem’ and *parèixer* ‘to seem’ incorporate modal information.

a) Regarding the verb *ser*, we can distinguish between two types of copular sentences: equative and descriptive sentences. In equative copular sentences, the predicate is a definite NP (or nominal clause): *La directora és la Maria*, *La Maria és la directora* ‘Mary is the director’. The verb number tends to be plural when one of the two elements is a plural NP: *La meva herència són aquests llibres*, *Aquests llibres són la meva herència* ‘My inheritance are these books’ (Dols/Mansell 2017, 176–179; Hualde 1992, 73–78; Ramos 2002, 1953–1992; Wheeler/Yates/Dols 1999, 521–527). In descriptive copular sentences, the complement can be an AdjP (*Ella és molt simpàtica* ‘She is very kind’), a PP (*Ella és de Girona* ‘She is from Girona’), an AdvP (*Ella és aquí* ‘She is here’) and a NP (*Ella és mestra de primària* ‘She is a school teacher’).

b) *Estar*, *quedar-se* and *romandre* ‘to remain’, and other aspectual quasi-copular verbs, make the duration of the state expressed by the complement explicit, while *resultar* and *esdevenir* express the achievement of a state of affairs: *En Pere està malalt* ‘Peter is ill’; *En Pere es va quedar cec* ‘Peter remained blind’; *En Pere va resultar lesionat* ‘Peter became injured’. On aspectual quasi-copular verbs and the quasi-copular behavior of some verbs of movement (*anar* ‘to go’, *venir* ‘to come’, *sortir* ‘to go out, to turn out’, etc.: *Tot sortirà molt bo* ‘lit. Everything will go out very good’, see Ramos (2002, 1978–1986), and GIEC (2016, 864–866).

c) The verbs *semblar* and *parèixer* ‘to seem, to appear’ are synonymous but they exhibit dialectal variation. They modalize the predicate expressed by the complement (a NP, an AdjP, a PP). Let’s compare *Aquest llibre és interessant* ‘This book is interesting’ with *Aquest llibre sembla interessant* ‘This book seems interesting’: the last sentence implies that it is not certain whether the book is interesting. On the properties of these copular verbs, see Ramos (2002, 1970–1972), and GIEC (2016, 863–864).

d) Descriptive copular sentences can involve both of the verbs *ser* and *estar*. For the sake of simplicity, we can say that *ser* is used when the property expressed by the complement is inherent (*La noia és bruna* ‘The girl is brunette’), and *estar* is used when the property is circumstantial (*La noia està trista* ‘The girl is sad’; *Estic de baixa* ‘I am on sick leave’). Nevertheless, the use of one or the other verb can be determined by the syntactic category of the complement (e.g. an AdvP of manner goes with *estar*: *Això està bé* ‘This is well’), and by the animate or non-animate nature of the subject.

Adjectives applied to an animate subject that describe the achievement of a state are selected by *estar*: *L'avi aviat estarà bo* 'The grandfather will be healthy soon'. However, there are some adjectives (and participles) that match with both *ser* and *estar*: *Els meus amics són animats* (an inherent property) 'My friends are cheerful'; *Els meus amics estan animats* (a non-inherent property) 'My friends are in a good mood'. Adjectives such as *orgullós* 'proud' or *gelós* 'jealous' only match with *estar* if they have a complement: *Sou orgullós* 'You are proud'; *Esteu orgullós del vostre fill* 'You are proud of your son'. For a list of predicative expressions with *estar*, see GIEC (2016, 867–871), Dols/Mansell (2017, 177–179).

e) If the subject is not animate, the choice of *ser* or *estar* is not clearly defined and depends on the dialect: *Aquest clima és humit* (an inherent property)/*Això està humit* (a circumstantial property) 'This climate is humid/This is humid', *El cafè és fred* (a description)/*El cafè està fred* 'The coffee is cold/The coffee has got cold'. The situation is similar with other adjectives (*buit* 'empty', *fluix* 'weak', *net* 'clean', etc.).

f) With participles (*cobert* 'covered', *escrit* 'written', *espatllat* 'broken', *mullat* 'wet', *ennuolat* 'overcast', *trenat* 'broken', *situat* 'situated', etc.), the verb is usually *estar*: *La carta ja està escrita* 'The letter is already written'; *El cel està ennuolat* 'The sky is overcast'; *Cadaqués està (situat) a la Costa Brava* 'Cadaqués is on the Costa Brava'.

g) With a locative PP or AdvP, the prototypical copular verb is *ser*: *Jo sóc a casa* 'I am at home'; *El teu cotxe és allà/al garatge* 'Your car is there/in the garage'. *Estar-se* and *estar* with a locative complement mean 'to stay, to remain': *L'Aina i la Gal·la s'estan a Olinda* 'Aina and Gal·la are staying in Olinda'; *La Joana (s')està de mestra a Camós/La Joana (s')està a Camós de mestra* 'Joan is (working) as a teacher in Camós'. The use of *estar* with a locative complement is spreading to some dialects and in urban areas: *Girona és/està a meitat de camí del mar i de la muntanya* 'Girona is half way between the sea and the mountain' (Hualde 1992, 78–79; Ramos 2002, 1994–2005; GIEC 2016, 871–876).

Temporal locations match with *ser*: *Ara és un quart de sis* 'It's a quarter past five'; *Avui és diumenge* 'It's Sunday'. However, *estar* can appear if the complement is a PP: *Avui és 3 de juliol/Avui estem a 3 de juliol* 'Today is July 3rd'.

h) Besides *ser* and *estar*, Catalan has the impersonal presentational or locative verb *haver-hi* (to *have* + locative clitic *hi*). The complementarity between *ser* and *haver-hi* is evident in *Els llibres eren damunt la taula* and *Damunt la taula hi havia els llibres* 'The books were on the table'. In the first sentence, *els llibres* provides known information while, in the second sentence, this NP provides new information. However, because of its impersonal nature, *haver-hi* cannot select a personal pronoun, in contrast with the verb *ser*: *Jo era allà* or *Hi era jo* 'I was there'. However, *haver-hi*, as opposed to *ser*, can select a bare NP: *Aquí hi ha pols* 'There is some dust here'; *Hi havia cotxes* 'There were some cars'. On the other hand, both *haver-hi* and *estar* can be used with a PP (*a la venda* 'for sale', *a l'abast* 'at reach', *en joc* 'at stake'): *Hi ha/Està a la venda aquesta casa* 'This house is for sale'.

The system described here is the most general and that recommended by the prescriptive grammar. However, the use of *estar* is increasing in urban areas due to Castilian influence to the detriment of *haver-hi* with a definite NP (*Hi ha en Biel?* ‘Is Biel here/there?’) and the use of *ser*: *Hi és, en Biel?* ‘Is Biel here?’ (GIEC 2016, 872–874).

9 Tenses

The verbal properties of tense, aspect and mood are expressed by means of inflection or highly grammaticalized periphrases (for more information ↗5.1 Word Classes, Inflectional Categories and Paradigms). Tense is a deictic feature, which locates the event or situation expressed by the verb temporally in relation to either the speech point or another point of reference. Aspect is a property that highlights how the event or situation is viewed, indicating whether the situation is presented globally, is in the process of developing, etc. Mood is a grammatical property related to sentence modality and to the assertive or non-assertive nature of the propositional content. In this section, temporal and aspectual properties of verbal tenses and of non-finite forms are analyzed. The opposition between the indicative and the subjunctive moods appears mainly in subordination and is analyzed in ↗5.3 The Complex Sentence.

9.1 Indicative

Tenses of the indicative mood offer clearer temporal and aspectual distinctions than those in the subjunctive. With regard to time, they may denote that the situation expressed by the verb is anterior, simultaneous or posterior either to the speech point or another point of reference in the past (Pérez Saldanya 2002, 2574–2576; GIEC 2016, 907–925). In Table 5, the basic temporal value of indicative tenses, exemplified by the third person singular of *parlar* (‘to speak, to talk’), are shown. Tenses oriented to the speech point are exemplified in (30a), and to a previous point of reference in (30b).

Table 5: Temporal distinctions in the indicative.

	ANTERIOR	SIMULTANEOUS	POSTERIOR
SPEECH POINT	simple and periphrastic past <i>parlà/va parlar</i> ‘(s)he spoke’	present <i>parla</i> ‘(s)he speaks’	future <i>parlarà</i> ‘(s)he will speak’
PREVIOUS TO THE SPEECH POINT	pluperfect <i>havia parlat</i> ‘(s)he had spoken’	imperfect <i>parlava</i> ‘(s)he spoke’	conditional <i>parlaria</i> ‘(s)he would speak’

- (30) a. Ahir *vaig fer* (periphrastic past) jo el sopar, ara el *prepara* (present) el pare i demà el *farem* (future) tu i jo.
 ‘I *cooked* dinner yesterday, our father *is cooking* it now and you and I *will do* it tomorrow’
- b. Va dir que ahir *havia fet* (pluperfect) ell el sopar, que en aquell moment el *preparava* (imperfect) el seu pare i que l’endemà el *farien* (conditional) ells.
 ‘He said that he *had cooked* dinner yesterday, their father *was cooking* it then and they *would do* it the next day’

As for aspect, verbal tenses may be imperfective, perfective and/or perfect. Tenses that indicate simultaneity are imperfective (present and imperfect). The simple and periphrastic past are perfective and all the compound tenses with the auxiliary *haver* ‘to have’ are perfect and in some cases also perfective: the present perfect (*ha parlat* ‘(s)he has spoken’), the pluperfect (*havia parlat* ‘(s)he had spoken’), the past anterior (*hagué parlat* o *va haver parlat* ‘(s)he had spoken’), the future perfect (*haurà parlat* ‘(s)he will have spoken’) and the conditional perfect (*hauria parlat* ‘(s)he would have spoken’). Those tenses posterior to the speech point are not marked aspectually (simple future and conditional simple).

a) As imperfective tenses, the present and the imperfect may have a progressive (31), habitual (32) or continuous (33) meaning.

- (31) a. Parla baix, que el nen *dorm*. (present)
 ‘Speak softly, the baby *is sleeping*’
- b. En aquell moment el nen *dormia*. (imperfect)
 ‘The baby *was sleeping* then’
- (32) a. Sempre *arribes* tard. (present)
 ‘You always *arrive* late’
- b. Sempre *arribaves* tard. (imperfect)
 ‘You always *arrived* late’
- (33) a. Fa quatre dies que *plou*. (present)
 ‘It *has been raining* for four days’
- b. Feia quatre dies que *plovia* (imperfect)
 ‘It *had been raining* for four days’

b) As past perfective tenses, the periphrastic and simple past are clearly distinct from the imperfect and the present perfect. Unlike the imperfect, they denote that the past situation expressed by the verb is conceived as a delimited whole and, if the opposite is not implied, takes place just once. Thus, in (34a), the past (*van tocar* ‘stroke’) refers to the moment when the clock started to strike, whereas the imperfect (*tocaven* ‘was striking’) highlights the ongoing process. In (34b), the past (*vas arribar* ‘you arrived’) denotes that (s)he arrived late just once, but the imperfect (*arribaves* ‘you were (usually) arriving’) implies that (s)he was usually arriving late.

- (34) a. En aquell just moment {*van tocar/tocaven*} les dotze al campanar.
 ‘The clock tower {*stroke/was striking*} twelve just then’
 b. L’any passat {*vas arribar/arribaves*} tard.
 ‘You {*arrived/were (usually) arriving*} late last year’

c) Furthermore, the past stands in contrast with the present perfect. Both tenses denote that the event or situation is anterior to the speech point. However, the relation that they establish with that time is different. The past is prehodiernal, meaning that it denotes that the situation occurred prior to today and it is not located within any interval that includes the speech point (35a). Conversely, the present perfect may express, among others, the hodiernal past meaning (35b), a past time within today or within any time period that includes the time of speech.

- (35) a. Se’n va anar {ahir/fa temps/l’any passat/*aquest matí/*avui/*aquesta setmana/*enguany}.
 ‘(S)he left {yesterday/some time ago/last year/*this morning/*today/*this week/*this year}’
 b. Se n’ha anat {*ahir/*fa temps/*l’any passat/aquest matí/avui/aquesta setmana/enguany}.
 ‘(S)he left {*yesterday/*some time ago/*last year/this morning/today/this week/this year}’

d) As a perfect tense, the present perfect may also exhibit different meanings that link a past situation to the speech point (Pérez Saldanya 2002, 2587–2592), specifically, a resultative (36a), experiential (36b), indefinite (36c) or inclusive meaning (36d), or that of informative relevance (36e).

- (36) a. Ja *hem parat* taula.
 ‘We *have already* set the table’ (→ ‘the table is set’)
 b. Pregunta-ho a la Marta que *ha estudiat* medicina.
 ‘Ask (it to) Marta, as she *has studied* medicine’ (→ ‘she has knowledge of medicine and thus she can help you’)
 c. Nosaltres també *hem estat* a Berlin.
 ‘We *have also been* to Berlin’ (→ ‘at least once before now’)
 d. Els meus pares sempre *han viscut* en aquesta casa.
 ‘My parents *have always lived* in this house’ (→ ‘and they are still living here’)
 e. *Ha tornat* a guanyar les eleccions!
 ‘(S)he *has won* the election again’

e) The pluperfect may denote a perfective or perfect meaning depending on the context, and thus it may work as an earlier past (37a) or as a past perfect (37b). In the

first case, the activity of 'leaving' takes place at 10 and is anterior to that expressed by the main clause; in the second one, it takes place before 10 and, as a result, there is no one concerned.

(37) a. Em va dir que se n'havien anat de casa a les 10.

'(S)he said that they *had left* home at 10'

b. A les deu ja se n'havien anat de casa.

'They *had already left* home at 10'

f) The other compound tenses exhibit more restrictive uses. The future perfect (38a) and the conditional perfect (38b) denote resultative meanings in addition to the temporal one of the parallel simple form. The past anterior is infrequent and only used in temporal subordinate clauses with connectives expressing that the event or the situation is immediately anterior to the past event expressed in the main clause (38c).

(38) a. Quan arribarem, ja se n'haurà anat tothom.

'When we arrive, everybody *will have already left*'

b. Em va dir que quan arribariem ja se n'hauria anat tothom.

'(S)he told me that everybody *would have already left* when we arrived'

c. Una vegada que {hagué/va haver} acabat de menjar, se n'anà.

'Once (s)he *had finished* eating, (s)he left'

g) Besides the basic uses just described, verbal tenses may express stylistic or modal values that alter their basic time or aspectual meanings. The most obvious case is that of the conditional. This tense may have the temporal meaning of future in the past, but also the modal meaning of unreality in the present or future. The temporal meaning is more or less linked to reported speech (39a) and the modal one occurs in the apodosis of conditional irrealis (39b) or in those sentences that, despite not being conditional, express an implicit condition related to the present (39c).

(39) a. Em va dir que tornaria l'endemà.

'(S)he told me that (s)he *would go* back the next day'

b. Si em fessis cas, no patiries tant.

'If you followed my advice, you *wouldn't suffer* so much'

c. Me n'aniria amb vosaltres al teatre, però no puc.

'I *would go* to the theatre with you, but I can't'

h) The future and the other prospective tenses may also be used in contexts where the idea of posterior time is blurred and various pragmatic connotations linked to the speech act emerge. Thus, the future may express obviousness (40a), courtesy (40b), surprise or reproach (40c) or uncertainty (40d). In some dialects, it may also express that something is inferred to be the case (40e), but, with this meaning, the periphrasis

deure + infinitive ‘should/may + infinitive’ is more common (40f) (75.3 The Complex Sentence).

- (40)a. Et *confessaré* que jo també he votat a favor.
‘I *will confess* that I also voted for it’
- b. Doncs, jo *voldré* una amanida.
‘So, I *will have* a salad’
- c. Ara em *vindràs* amb això?
‘What kind of excuse is this?’ (lit. ‘Now *will you come* with that?’)
- d. Que *serà* això que porta?
‘What *will* that thing (s)he is carrying *be*?’
- e. Ja telefonarem més tard que ara *estaran* sopant.
‘We will phone later, as they *should be having* dinner now’
- f. Ja telefonarem més tard que ara *deuen* estar sopant.
‘We will phone later, as they *should be having* dinner now’

i) Being the less marked tense in the verbal system, the present may refer to expected or planned future situations (41a) or, with a past meaning, may be used in the so-called historic present (41b) or in the narrative past, alternating with the past tense to make the narration more vivid (41c).

- (41) a. Demà *presento* la dimissió i *deixo* l’empresa.
‘I *resign* tomorrow and leave the company’
- b. Carles Aribau *escriu* l’*Oda a la Pàtria* l’any 1832.
‘Carles Aribau *writes* the *Ode to Motherland* in 1832’
- c. L’altre dia passejava per l’albereda quan de sobte *sento* que algú em *crida*.
‘I was walking along the boulevard some days ago, when suddenly I *hear* someone *calling* me’

j) Not being marked in the past sphere, the imperfect may also be used instead of the conditional to express situations posterior to a past moment (42a), and instead of the past in the so-called narrative imperfect (42b).

- (42) a. Ens va confessar que l’endemà *presentava* la dimissió.
‘(S)he admitted that (s)he *resigned* the next day’
- b. Aquell mateix any *esclatava* la guerra.
‘The war *broke* out that very year’

9.2 Subjunctive

As a marked mood (75.3 The Complex Sentence), the subjunctive includes fewer verbal tenses than the indicative and a tense in the subjunctive may be related in meaning to two or more tenses in the indicative, as shown in Table 6.

Table 6: Indicative and subjunctive tenses.

INDICATIVE	SUBJUNCTIVE
Present (<i>parla</i> '(s)he speaks')	Present (<i>parli</i> '(s)he speaks')
Future (<i>parlarà</i> '(s)he will speak')	
Present perfect (<i>ha parlat</i> '(s)he has spoken')	Present perfect (<i>hagi parlat</i> '(s)he has spoken')
Future perfect (<i>haurà parlat</i> '(s)he will have spoken')	
Periphrastic/simple past (<i>va parlar/parlà</i> , '(s)he spoke')	Imperfect (<i>parlés</i> '(s)he spoke')
Imperfect (<i>parlava</i> '(s)he spoke')	
Conditional (<i>parlaria</i> '(s)he would speak')	
Pluperfect (<i>havia parlat</i> '(s)he had spoken')	Pluperfect (<i>hagués parlat</i> '(s)he had spoken')
Past anterior (<i>hagué parlat/va haver parlat</i> '(s)he had spoken')	
Conditional perfect (<i>hauria parlat</i> '(s)he would have spoken')	

In contrast with the indicative, the subjunctive has no prospective tenses and the same tense may express, depending on the context, either a simultaneous or a posterior process, as in (43a) with the present subjunctive. There does exist a past perfective subjunctive tense (*vagi parlar*), built in similar way to the periphrastic past indicative. However, this is a rather odd form and, to express this meaning, the imperfect subjunctive is normally used in the standard language, as shown in (43b), in instances where the periphrastic past subjunctive (*vagi oposar*) and the imperfect subjunctive (*oposés*) may be alternatives. Besides functioning as a past perfective, the imperfect subjunctive may be used with the meaning of a past imperfective (43c), of a future in the past (43d) and with different modal values related to the sphere of the present. For instance, it may express present or future in the protasis of irrealis conditionals (43e) or in optative clauses (43f), or with a polite imperative meaning in negative sentences (43g).

- (43) a. És possible que {en aquest moment/quan arribem demà} *estigui treballant*.
 'It is possible that (s)he {is working now/will be working when we arrive tomorrow}'
 b. No crec que s'hi {*vagi oposar/oposés*}.
 'I don't think (s)he is opposed to it'

- c. Em sabia greu que es *trobés* malament.
'I was sorry that (s)he *felt* bad'
- d. Volia que *tornéssim* l'endemà.
'(S)he would rather that we *came back* the next day'
- e. Si *tingués* diners, invertiria en aquesta empresa.
'If I *had* money, I would invest in this company'
- f. Tant de bo es *trobés* ara aquí!
'I wish (s)he *were* here now'
- g. No us *penséssiu*, però, que hi estic interessat!
'However, you *should not think* that I'm interested in it'

10 Verbal periphrases

A verbal periphrasis consists of an auxiliary verb, which can afford either an aspectual or a modal meaning, and a main verb in a non-finite form, which constitutes the predicator. Thus, in *Pot ploure* 'It may rain', the auxiliary *poder* 'may' and the main verb in infinitive *ploure* 'rain' constitute a verbal periphrasis, in which the predicator is expressed by the main verb and the auxiliary provides the modal meaning of possibility (Hernanz/Rigau 1984; Gavarró/Laca 2002, 2667–2673; GIEC 2016, 945–948).

10.1 Aspectual periphrases

Aspectual periphrases show in which stage of accomplishment the process expressed by the main verb is, or highlight whether it is an iterative situation or process, if it has been interrupted, etc. (Gavarró/Laca 2002, 2683–2710; GIEC 2016, 952–957). Regarding the aspectual meaning expressed, the main verb may occur in infinitive, in gerund or in participle as shown in Table 7, in which the most usual periphrases are summarized and exemplified.

Table 7: Aspectual periphrases.

PERIPHRASES	MEANING	EXAMPLES
<i>començar a</i> + inf	beginning	<i>Hem començat a llegir</i> les cartes. 'We have started to read the letters'
<i>acabar de</i> + inf	completion	<i>Hem acabat de llegir</i> les cartes. 'We have finished reading the letters'
	immediate past	<i>Se n'acaben d'anar</i> . 'They just left'

PERIPHRASES	MEANING	EXAMPLES
<i>estar</i> + ger	progression	No entris, que <i>està dormint</i> . 'Don't go in, (s)he <i>is sleeping</i> '
<i>continuar</i> + ger	ongoing	<i>Continua vivint</i> al mateix lloc de sempre. '(S)he <i>is still living</i> in the same place as ever'
<i>anar</i> + ger	internal process	El cel <i>anava fent-se</i> cada cop més negre. 'The sky <i>became</i> darker and darker'
<i>tornar a</i> + inf	repetition	Ara <i>torna a fumar</i> després de tants anys. '(S)he <i>smokes again</i> after so many years'
<i>soler</i> + inf	habitual iteration	Els diumenges <i>solen anar</i> d'excursió. 'They <i>usually go</i> hiking on Sundays'
<i>tenir</i> + part	result	Ja <i>tinc redactat</i> el capítol. 'I <i>have already written</i> the chapter'

Unlike in other Romance languages, the periphrasis *anar a* + infinitive does not have a future meaning with the auxiliary verb in present; the simple future is used instead with this meaning: *Demà donarà una conferència* '(S)he will give a speech tomorrow' (and not **Demà va a donar una conferència* '(S)he is going to give a speech tomorrow'). This periphrasis expresses imminence and is used mainly to highlight that a situation that was about to take place has been stopped: *Anava a sortir, però va començar a ploure* 'I was going to leave, but it started to rain'.

10.2 Modal periphrases

In terms of their meaning, modal periphrases may be either root or epistemic. Root modal periphrases are agent-oriented (usually, the subject of the sentence); they express the agent's 'ability', 'permission', 'obligation', etc. to do something. Conversely, epistemic modal periphrases are sentence-oriented and indicate whether the state of affairs is necessary, possible or probable, according to the speaker's point of view (Gavarró/Laca 2002, 2711–2716; GIEC 2016, 950–952).

a) The periphrasis *poder* + infinitive is used with a root meaning to express ability (*Pots fer tot el que et proposis* 'You can do whatever you intend') or permission (*Pots anar-te'n quan vulguis* 'You can/may go whenever you like'). It is used with epistemic meaning to express probability (*Demà pot nevar per sota dels cinc cents metres* 'It may snow below five hundred meters tomorrow').

b) The periphrasis *haver de* + infinitive is mainly used with the root meaning of obligation (*Hem d'esforçar-nos més* 'We have to/must try hard to do it'). However, it may also be used with an epistemic meaning to highlight that something is inferred to be the case, similar to 'must/should' in English: *Ha hagut de ser ell, perquè no ens mira*

als ulls ('He must/should have done it, as he doesn't look us in the eyes'). This periphrasis is also used with a future meaning.

c) Conversely, the periphrasis *deure* + infinitive is used more or less with this inferable epistemic meaning (*Encara deu estar dolgut amb nosaltres, perquè no s'ha acostat a saludar-nos* 'He may still be grieved with us, since he hasn't come to greet us'). This periphrasis is obsolete in most dialects to express obligation: *Deus obeir els pares* 'You must obey your parents'.

d) The full verbs *saber* 'know' and *voler* 'want' may also exhibit meanings similar to modal auxiliaries when combined with an infinitive form and thus expressing, respectively, the ability to do something (*Ella també sap escalar en gel* 'She also knows how to ice climb') and the intention to do it (*Vol anar al cinema* '(S)he intends to go to the cinema') or the probability that something takes place immediately (*Sembla que vol ploure* 'It seems like it's about to rain').

e) The periphrasis *gosar* + infinitive also has a root meaning equivalent to 'have the courage to do something', as in *No gosaven dir res* ('They didn't dare to say anything').

11 Affirmative and negative polarity

Affirmative and negative are two poles on a binary scale: the positive pole, which in declarative sentences is identified with the truth of the propositional content, and the negative pole, identified with the falsehood of the propositional content.

a) The affirmative polarity is the unmarked one, and thus affirmative clauses do not generally display any mark of polarity (44a). However, they may be introduced by the adverb *sí*, which is always followed by the conjunction *que*, to affirmatively emphasize their propositional content (44b) or some of their constituents (44c). In the latter context, the focalized constituent *demà* 'tomorrow' occurs in the left periphery.

- (44)a. *Demà hi ha classe.*
'There are lessons tomorrow'
- b. *Sí que hi ha classe demà.*
'There are lessons tomorrow'
- c. *Demà, sí que hi ha classe.*
'Tomorrow, there are lessons'

The prototypical device for expressing the negative polarity is the adverb *no*, which occurs before the verb or the verb phrase with a proclitic pronoun (45a). The adverb *no* may be reinforced by the adverb *pas*, which adds a contrastive meaning to negation (Espinal 2002, 2748–2752; GIEC 2016, 1309–1310), as in (45b), where a presupposition is negated, or in (45c), in which it adds a menacing or warning tincture to a prohibition.

- (45) a. Demà *no* hi ha classe.
 ‘There are *no* lessons tomorrow’
 b. –Us ha agradat el teatre? –*No* hem anat pas al teatre.
 ‘–Did you enjoy the play? –We *did not* go to the theatre’
 c. *No* diguis pas això, eh?
 ‘*Don’t* say that, ok?’

b) The adverbs *sí* and *no* are also used to answer a total interrogative sentence (46a) or as a reply or reinforcement of what the interlocutor has previously said (46b). In those contexts, a sentence-constituent may occur left-dislocated, before the adverbs (46c); these may also be followed by a clause containing the very same adverbs, which strengthen the answer or the reply (46d).

- (46) a. –Demà hi ha classe? –*Sí/–No*.
 ‘–Are there lessons tomorrow? –*Yes/–No*’
 b. –Demà hi ha classe. –*Sí/–No*.
 ‘–There are lessons tomorrow. –*Yes/–No*’
 c. Demà, *sí/no*.
 ‘Tomorrow, *yes/no*’
 d. *Sí, sí* que hi anirem; *No, no* hi anirem.
 ‘*Yes, we WILL* go there; *No, we WON’T* go there’

The negative particle may also occur before the strictly negated constituents in the so-called constituent-negation: *Vaig enviar la nota de protesta no (pas) al director, sinó a la secretària* ‘I DIDN’T send the protest note to the director, but to the secretary’.

c) In negative clauses, the adverb of negation may occur together with other negative words, known as negative polarity items (NPI). Specifically, they are negative quantifiers (47a), the negative adverb *tampoc* ‘neither/either’ (47b) or coordinate phrases with the conjunction *ni* ‘neither/nor, (not) either/or’ (47c). Between the adverb *no* and the NPI, a relation of negative concord is established and thus negation affects various constituents throughout the sentence. The adverb of negation is obligatory when the NPI occurs after the verb (47a–c). It is, however, optional when the NPI occurs before the verb (47d).

- (47) a. Sobre aquest tema, no he dit *mai* res a *ningú*.
 ‘About this subject, I have *never* said *anything* to *anybody*’
 b. Jo no podré assistir-hi *tampoc*.
 ‘I won’t be able to attend *either*’
 c. No he pogut comprar *ni* el pa *ni* els ous.
 ‘I couldn’t buy *either* the bread *or* the eggs’
 d. *Mai (no)* he dit *res* a *ningú* sobre aquest tema.
 ‘I have *never* said *anything* to *anybody* on this subject’

Apart from the adverb *no*, the negative concord relation may be established with other lexically negative items, such as the preposition *sense* ‘without’ (*sense l’ajuda de ningú* ‘without anybody’s help’) or such verbs as *impedir* ‘forbid, prevent, block, obstruct’ or *negar* ‘deny’: *Ens van impedir que anéssim enlloc* ‘They forbade us to go anywhere’.

d) Negative quantifiers may also be used in non-negative clauses with a non-specific meaning equivalent to ‘anything, anywhere...’. However, this usage is only possible in contexts where a positive state of affairs can be put into relation, at least implicitly, to a negative one (Espinal 2007, 51–52; GIEC 2016, 1306–1309). This is the case, for instance, in the protasis of conditionals (*Si necessites res, demana-m’ho* ‘If you need *something/anything*, ask for it’), in total interrogatives (*L’has vist enlloc?* ‘Have you seen her/him *anywhere?*’), with verbs expressing ‘fear’, ‘doubt’, ‘opposition’, ‘obstacle/hindrance’, etc. (*Dubto que t’ho hagi demanat mai* ‘I doubt that I have ever asked for it’) or in the codas of inequality comparatives (*Treballa més que ningú* ‘(S)he works harder than *anybody*’).

e) For the same reason, some of these contexts also allow expletive negation (Espinal 2002, 2776–2779; GIEC 2016, 1313–1314). This is the case, for instance, with predicates expressing ‘fear’ (*Tinc por que no faci una animalada* ‘I’m afraid that (s)he may commit an atrocity’), with codas of inequality comparatives (*Estudia més que no (pas) tu* ‘(S)he studies more than you’) or with subordinate clauses headed by *abans* ‘before’ (*Explica-ho tu abans (que) no t’ho preguntin ells* ‘Explain it before they don’t ask about it’), by *fins* ‘until’ when the main clause is negative (*No ens ajudaran fins que no els ho demanem* ‘They won’t help us until we ask for it’) and with *de poc* or *per poc* ‘almost, nearly’, lit. ‘for little’ (*Per poc que no tenim un accident* ‘We almost/nearly had an accident’).

12 Bibliography

- Alsina, Àlex (2016), *Catalan*, in: Adam Ledgeway/Martin Maiden (edd.), *The Oxford Guide to the Romance Languages*, Oxford, Oxford University Press, 363–381.
- Bartra Kaufmann, Anna (2002), *La passiva i les construccions que s’hi relacionen*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 2111–2179.
- Bartra Kaufmann, Anna/Suñer, Avel·lina (1992), *Functional Projections Meet Adverbs*, Catalan Working Papers in Linguistics 1, 45–85.
- Bel, Aurora (2002), *Les funcions sintàctiques* in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 1075–1147.
- Brucart, Josep M. (2002), *Els determinants*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 1435–1516.
- Brucart, Josep M./Rigau, Gemma (2002), *La quantificació*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 1517–1589.
- Dols, Nicolau/Mansell, Richard (2017), *Catalan. An Essential Grammar*, London/New York, Routledge.
- Espinal, M. Teresa (2002), *La negació*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 3, Barcelona, Empúries, 2727–2797.

- Espinal, M. Teresa (2007), *Licensing Expletive Negation and Negative Concord in Romance Languages*, in: Franck Floricic (ed.), *La négation dans les langues romanes*, Amsterdam/Philadelphia, Benjamins, 49–74.
- Fabra, Pompeu (1956), *Gramàtica catalana*, Barcelona, Teide.
- Gavarró, Anna/Laca, Brenda (2002), *Les perífrasis temporals, aspectuals i modals*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 3, Barcelona, Empúries, 2663–2726.
- Gràcia i Solé, Lluïsa (1989), *Els verbs ergatius en català*, Ciutadella, Institut Menorquí d'Estudis.
- Hernanz, M. Lluïsa (2002), *L'oració*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 993–1073.
- Hernanz, M. Lluïsa/Rigau, Gemma (1984), *Auxiliaritat i reestructuració*, *Els Marges* 31, 29–51.
- Hualde, José Ignacio (1992), *Catalan*, London/New York, Routledge.
- GIEC = Institut d'Estudis Catalans (2016), *Gramàtica de la llengua catalana*, Barcelona, Institut d'Estudis Catalans.
- Martí i Girbau, Núria (2002), *El SN: els noms*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 1281–1335.
- Nogué, Neus (2015), *Catalan*, in: Konstanze Jungbluth/Federica Da Milano (edd.), *Manual of Deixis in Romance Languages*, Berlin/Boston, De Gruyter, 206–239.
- Payrató, Lluís (2002), *L'enunciació i la modalitat oracional*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 1149–1220.
- Pérez Saldanya, Manuel (2002), *Les relacions temporals i aspectuals*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 3, Barcelona, Empúries, 2567–2662.
- Pérez Saldanya, Manuel (2015), *Paradigms as Triggers of Semantic Change: Demonstrative Adverbs in Catalan and Spanish*, *Catalan Journal of Linguistics* 14, 113–135.
- Pérez Saldanya, Manuel/Rigau, Gemma (2007), *Els adverbis de lloc intransitius i la formació de construccions amb adverbis postposats*, *Estudis Romànics* 29, 61–80.
- Picallo, M. Carme (2002), *L'adjectiu i el sintagma adjectival*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 1641–1688.
- Ramos, Joan-Rafael (2002), *El SV II: La predicació no verbal obligatòria*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 1951–2044.
- Rigau, Gemma (1986), *Some Remarks on the Nature of Strong Pronouns in Null-Subject Languages*, in: Ivonne Bordelois/Heles Contreras/Karen Zagona (edd.), *Generative Studies in Spanish Syntax*, Dordrecht, Foris, 143–164.
- Rigau, Gemma (1988), *Strong Pronouns*, *Linguistic Inquiry* 19/3, 503–510.
- Rigau, Gemma (1997), *Locative Sentences and Related Constructions in Catalan: ésser/haver Alternation*, in: Amaya Mendikoetxea/Miriam Uribe-Etxebarria (edd.), *Theoretical Issues at the Morphology-Syntax Interface*, Bilbao, Euskal Herriko Unibersitatea, 395–421.
- Rigau, Gemma (2002), *Els complements adjunts*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 2054–2110.
- Rigau, Gemma (2013), *La preposició silent d'alguns verbs de moviment local*, *Els Marges* 100, 125–132.
- Rosselló, Joana (2002), *Verb i arguments verbals*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 1857–1949.
- Sancho Cremades, Pelegrí (2002), *La preposició i el sintagma preposicional*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 1689–1796.
- Solà, Joan (1973), *Estudis de sintaxi catalana/2*, Barcelona, Edicions 62.
- Solà, Joan (1987), *Qüestions controvertides de sintaxi catalana*, Barcelona, Edicions 62.
- Solà, Joan (1994), *Sintaxi normativa. Estat de la qüestió*, Barcelona, Empúries.
- Solà, Joan (2009), *Les gramàtiques de Pompeu Fabra de 1956 i 1946*, in: Jordi Mir/Joa Solà (edd.), *Pompeu Fabra. Obres completes*, vol. 6, Barcelona/València/Palma (Mallorca), Proa/Edicions 3i4/Moll, 273–458.

- Todolí, Júlia (2002), *Els pronoms*, in: Joan Solà et al. (edd.) (2002), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 1337–1433.
- Wheeler, Max W./Yates, Alan/Dols, Nicolau (1999), *Catalan: A Comprehensive Grammar*, London/New York, Routledge.

5.3 The Complex Sentence

Abstract: This chapter deals with complex sentences and their different constructions. Whenever a construction may contain a dependent clause (or a complex phrase), every possibility is explained. Moreover, the conjunctions and, in general, the connective elements that define each complex construction are analyzed. The connective relations may be either coordination, where the related elements are functionally equivalent, or subordination, where a relation of hierarchical syntactical dependency is established between the clause introduced by the conjunction and another clause or between complex phrases. Section 1 is devoted to coordinating constructions and the following sections to subordinating ones, including the use of the indicative and the subjunctive moods in subordinate clauses. In section 2 content clauses are discussed; section 3 deals with relative clauses, 4 with the indicative and subjunctive mood in subordinate clauses, 5 with adverbial temporal constructions, 6 with comparatives and other clauses of degree, 7 with causal, purpose and result clauses, and 8 with conditional and concessive constructions.

Keywords: Catalan syntax, coordination, conjunctions, relative clauses, content clauses, comparative constructions, causal, purpose and result clauses, conditional and concessive clauses, mood

1 Coordination: Coordinating conjunctions

Coordinating constructions consist of two or more functionally equivalent members that are linked by conjunctions (*i* ‘and’, *ni* ‘nor’, *o* ‘or’, *però* ‘but’, etc.). The coordinate constituents may be clauses (*Tu m’ajudes a mi i jo t’ajudo a tu* ‘You help me and I help you’), phrases (*amb els meus germans o sense ells* ‘with my brothers or without them’) or phrasal constituents (*tots i cadascun dels inscrits* ‘each and every one registered’). In general, coordination of clitics is not possible or sounds unnatural, as in the case of the article (*els i les participants* ‘the (m.) and the (f.) participants’).

According to the relations expressed, coordinating conjunctions are classified as copulative, disjunctive and adversative. Copulative conjunctions express addition, disjunctive conjunctions alternation, and adversative conjunctions opposition. Coordinating conjunctions can be placed between the coordinate elements, as in the examples shown in the previous paragraph. However, in some contexts, they can also establish correlations and be placed before each coordinate constituent: *o tu o jo* ‘either you or I’; *ni tu ni jo* ‘neither you nor I’; *tant tu com jo* ‘both you and me’ (Bonet/Solà 1986, 314; Serra/Prunyonosa 2002; GIEC 2016, 965–967, 977–991).

<https://doi.org/10.1515/9783110450408-008>

a) The basic copulative conjunctions are *i* ‘and’ and *ni* ‘neither/nor’. The coordinator *i* may occur in affirmative contexts (*Vindrà demà i demà passat* ‘(S)he will come tomorrow and the day after tomorrow’), in negative ones (*No vindrà demà i tampoc demà passat* ‘(S)he will not come tomorrow nor the day after tomorrow’), or in a combination of both (*Vindrà demà i no demà passat* ‘(S)he will come tomorrow and not the day after tomorrow’). When more than two coordinate constituents occur, generally speaking, *i* is only used between the last two (*Vindrà demà, demà passat i l’altre* ‘(S)he will come tomorrow, the day after tomorrow and the day after that’). However, it may occur between each member, resulting in an effect of emphasis or insistence (*Vindrà demà i demà passat i l’altre* ‘(S)he will come tomorrow and the day after tomorrow and the day after that’).

b) Conversely, the conjunction *ni* ‘neither, nor’ only occurs in negative coordination (*No vindrà demà ni demà passat* ‘(S)he will not come tomorrow nor the day after tomorrow’) and may appear in correlations (*Ni vindrà demà ni demà passat* ‘(S)he will come neither tomorrow nor the day after tomorrow’).

c) In affirmative contexts, both the locution *així com* ‘as well as’ (1a) and the correlative conjunction *tant... com* ‘both... and’ (1b) may also be used.

- (1) a. *Vindrà demà així com demà passat.*
‘(S)he will come tomorrow *as well as* the day after tomorrow’
- b. *Vindrà tant demà com demà passat.*
‘(S)he will come *both* tomorrow *and* the day after tomorrow’

The additive value of copulative conjunctions may be strengthened by means of focusing adverbs expressing inclusion (*també* ‘also’, *fins i tot* ‘even’) or exclusion (*tampoc* ‘neither’):

- (2) a. *Vindrà demà i {també/fins i tot} demà passat.*
‘(S)he will come tomorrow *and* {*also/even*} the day after tomorrow’
- b. *No vindrà demà ni tampoc demà passat.*
‘(S)he will not come tomorrow *nor* the day after tomorrow’

d) The disjunctive conjunction *o* ‘or’ may occur in exclusive disjunctions, in which the members are presented as incompatible (3a), or in inclusive disjunctions, where the conjuncts are not incompatible (3b).

- (3) a. *Què vols per a sopar, carn o peix?*
‘What would you like for dinner, meat *or* fish?’
- b. *Podeu menjar o beure tot el que us abelleixi.*
‘You can eat *or* drink whatever you like’

In exclusive disjunctions, the conjunction *o* can be repeated before each coordinate member (*o carn o peix* ‘either meat or fish’) and can be reinforced with the particle *bé* ‘either’ (lit. ‘well’) (*o bé carn o bé peix* ‘either meat or fish’) or with *si no* ‘if not’ before the last conjunct (*carn o si no peix* ‘meat or, if not, fish’). In negative contexts, *ni* (lit. ‘neither’) is used instead of *o*: *sense carn ni peix* ‘without meat or fish’.

e) With a distributive meaning close to copulative and disjunctive coordination, there are a number of correlative structures with adverbs or other units repeated in initial position (*ara... ara* ‘now... now’, *adés... adés* ‘then... then’, *ara... adés* ‘now... then’, *aquí... aquí* ‘here... here’, *qui... qui* ‘someone... someone’, *sigui... sigui* ‘be... be’, etc.). These correlations designate pairs of alternating situations or situations considered as a whole: *Ara plou ara neva* ‘Now it’s raining, now it’s snowing’; *Ja plora ja riu*, ‘Now he’s crying, now he’s smiling’; *Tothom estava atent al que deien, qui dret qui assegut* ‘Everybody paid attention to what they were saying, some standing, some sitting’.

f) The basic adversative conjunctions are *però* and *sinó* (both meaning ‘but’ in English, but with different syntactic structures). The first one expresses a non-excluding (or partial) opposition and requires that one of the two conjuncts be negative (4a) or that some sort of contrast between them be established (4b) (Cuenca 1991). The second one, however, expresses an excluding (or total) opposition and requires that the first conjunct be negative (4c). If the second conjunct is a finite clause, it is usually preceded by *que* ‘that’ (4d).

- (4) a. És molt intel·ligent, *però* no té ganes d’estudiar.
 ‘(S)he is very clever, *but* does not feel like studying’
 b. Van arribar al cim exhausts *però* contents.
 ‘They reached the summit exhausted *but* happy’
 c. No m’ho ha dit el pare, *sinó* l’àvia.
 ‘It was not told by my father, *but* by my grandmother’
 d. No vingueren ells, *sinó que* anàrem nosaltres a casa seva.
 ‘They did not come, *but* we went to their house’

With a non-excluding value, there is also the locution *mentre que* ‘whereas’ ‘while’, with an originally temporal meaning; it coordinates clauses that establish a multiple contrast, as in (5), where the two people (*ella* ‘she’ and *mi* ‘I’, lit. ‘me’) and what each one prefers or likes are contrasted.

- (5) Ella prefereix vetllar, *mentre que* a mi m’agrada matinar.
 ‘She prefers staying up late, *whereas* I like getting up early’,

g) The coordinators *i*, *o* and *però* may also function as discourse connectives when they are in sentence-initial position and set a discourse-pragmatic link with the

preceding co-text (6a,b). Like in Old Catalan, *però* may work as a concessive connective adverb as well (6c).

- (6) a. Em va dir que ja no m'estimava. *I amb això, va acabar.*
 '(S)he told me that (s)he didn't love me anymore. *And*, with that, (s)he put an end to it'
- b. Els meus pares també van haver d'emigrar a França. *Però això ja t'ho explicaré un altre dia.*
 'My parents had also had to migrate to France. *But* I will explain it to you one day'
- c. Tingueu en compte, *però*, que demà farà mal oratge.
 'Take into account, *however*, that there will be bad weather tomorrow'

2 Content clauses

Subordinate content clauses (or noun clauses) are functionally equivalent to a NP and may have an infinitive (7a) or a finite verb, which can be either in the indicative (7c–f) or in the subjunctive (7b). According to the sentence modality, content clauses may be declarative (7a–c), interrogative (7d–e) or exclamatory (7f).

- (7) a. Necessito *demanar-te un favor*.
 'I need to ask you for a favor'
- b. Necessito *que em facis un favor*.
 'I need you to do me a favor'
- c. Diu *que et demanarà un favor*.
 '(S)he says that (s)he will ask for a favor'
- d. Em preguntà *si l'ajudariem*.
 '(S)he asked whether we'd help him/her'
- e. Demana a l'Anna *què vol i per a quan ho necessita*.
 'Ask Anna what she wants and when she needs it'
- f. No pots imaginar *que divertit que és ni quants acudits que conta*
 'You can't imagine how funny (s)he is nor how many jokes (s)he tells'

This section deals with declarative content clauses. Regarding interrogative and exclamatory content clauses, see ↗ 5.4 Modality and Information Structure: Focus, Dislocation, Interrogative and Exclamatory Sentences. As for the alternation of the indicative and the subjunctive moods in content clauses, see section 4 in this chapter.

a) Declarative clauses are introduced by the conjunction *que* 'that', if the verb is in a finite form (7b–c), and without a conjunction, if an infinitive occurs (7a). They can have any of the syntactic functions of a NP. They can be selected by a verbal predicate and act as the subject of the predication (*M'agrada que et trobis més bé* 'I am glad that

you feel better'), the direct object (*Sento que te n'hagis d'anar* 'I am sorry that you have to go'), the indirect object (*Donem molt de valor que ens hagis dit la veritat* 'We appreciate the fact that you've told the truth') or a prepositional object (*El govern insisteix a reforçar la seguretat* 'The government insists on strengthening security'). A declarative clause can also act as the complement of a preposition (*Parla sense mirar als ulls* '(S)he speaks without looking me/us/you in the eye'), of a noun (*l'esperança de veure'l bé* 'the hope of seeing him/her well'), of an adjective (*content de poder comptar amb vosaltres* 'happy to rely on you'), and sometimes of an adverb (*independentment que ho hagi dit el president* 'regardless of whether the president has said it'). Catalan never drops the conjunction *que* in front of finite declarative clauses.

b) As shown by the alternations of (8), the unstressed prepositions *de* 'of', *a* 'to', *en* 'in/on' and *amb* 'with' are dropped before *que* 'that' in finite clauses (8a), and *en* or *amb* can be changed into *a* or *de* before an infinitive clause (8b). For more details see 75.2 The Simple Sentence.

- (8) a. S'han alegrat {*de* parlar amb tu/Ø *que* parlessis amb ells}.
 'They were thrilled {to talk to you/Ø that you talked to them}'
 b. Confia *en* tu/Confia *a/de* rebre el teu ajut.
 '(S)he relies on you'/'(S)he relies on having your help'

c) In general, infinitive subordinate clauses have an elliptical subject. The subject is usually retrieved from its coreference with an argument of the main clause, which acts as the antecedent (Alsina 2002, 2397–2411; GIEC 2016, 1013–1015). If the main clause contains only one argument besides the subordinate clause, the subject is interpreted in relation to this argument, regardless of whether it is the subject (9a), the direct object (9b) or the indirect object (9c). If the main clause has more than one argument, the interpretation depends on the verb in the main clause. For instance, it is interpreted in relation to the subject in the cases of verbs like *prometre* 'to promise' (9d), and in relation to on the indirect object with verbs like *prohibir* 'to forbid' (9e).

- (9) a. Necessitem parlar amb tu.
 'We need to talk with you'
 b. La va animar saber que havia aprovat l'examen.
 'Knowing that she had passed the exam encouraged her'
 c. Em molesta haver de decidir tota sola.
 'Having to decide on my own annoys me'
 d. He promès a la Maria no dir res a ningú.
 'I've promised Mary not to tell anybody anything'
 e. El pare m'ha prohibit parlar d'això
 'My father has forbidden me to talk about this'

When the main clause does not contain any argument which could be the potential antecedent, the elliptical subject of the infinitive clause is interpreted as generic or arbitrary: *Protestar ara no serveix de res* 'Protesting now makes no sense'; *No es permet fumar* 'Smoking is not allowed'. Subordinate clauses depending on a noun may also receive this interpretation: *Analitzen la conveniència de substituir el sistema de rec* 'They analyze the convenience of replacing the irrigation system'. The same situation takes place when the clause depends on an adjective with a passive meaning such as *fàcil* 'easy' in *És un examen molt fàcil d'aprovar* 'It is an easy exam to pass'.

d) Predicates selecting a subordinate clause with the verb in the indicative mood do not usually accept it in the infinitive construction. However, alternation may occur if the clause depends on verbs like *jurar* 'to swear', *prometre* 'to promise' or *confessar* 'to confess' (10a), although the use of the indicative is more natural. Conversely, predicates selecting a clause in the subjunctive accept the infinitive too. In general, subordinate clauses either in the subjunctive or in the infinitive have a complementary distribution, as in (10b), where the subjunctive is used if the subject of the subordinate clause is different from that of the main clause, and the infinitive if both subjects are coreferent. However, sometimes, both subordinate structures may alternate, as in the case of the verb *negar* 'to deny' or other verbs expressing prohibition (10c).

- (10) a. Prometeren {que *acabarien* aviat/*acabar* aviat}.
'They promised {that they *would finish* soon/*to finish* soon}'
- b. Necessito {que *parlis amb el pare*/*parlar amb el pare*}.
'I need {you *to talk with our father*/*to talk with my father*}'
- c. M'han prohibit {que *digui/dir*} res.
'I've been forbidden *to say* anything'

e) Infinitive clauses functioning either as a subject or as a direct object may be preceded by the preposition *de* 'of' when following the main verb (Bonet 2002, 2377–2378; GIEC 2016, 1009–1011). Functioning as postverbal subjects, they accept the preposition if they depend on verbs with a psychological meaning (11a) and certain verbs expressing preference or obligation (11b). Functioning as direct objects, the preposition *de* is possible in the majority of cases (11c) and required with verbs like *dir* (with the meaning 'to suggest'), *pregar* 'to request' or *provar* 'to try to' (11d).

- (11) a. M'agrada (*de*) parlar amb vosaltres.
'I like talking to you'
- b. Et toca (*de*) dir què en penses.
'It is your turn to say what you think'
- c. Han aconseguit (*d'*)aprovar la reforma.
'They have got to approve the reform'
- d. M'han dit *d'*acompanyar-los.
'They've suggested that I go with them'

f) There are some verbs, such as *considerar* ‘to consider’ o *imaginar(-se)* ‘to imagine’, which accept a small clause, i.e. a subordinate clause which uses neither the conjunction *que* ‘that’ nor the copula *ser* ‘to be’. In (12a), for instance, the subordinate is a copulative clause introduced by *que*, whereas in (12b) it is a small clause that is semantically equivalent. In the last sentence, the subject of the small clause (*en Jaume*) acts as the direct object of the main clause, while the subject complement of the equivalent copular clause (*la persona més adequada* ‘the most suitable person’) is the predicate of the small clause.

(12) a. Considerem *que en Jaume és la persona més preparada*.

‘We consider *that Jaume is the most suitable person*’

b. Considerem *en Jaume la persona més adequada*.

‘We consider *Jaume the most suitable person*’

g) Causative (*fer* ‘to make’, *deixar* ‘to let’) and perception verbs (*sentir* ‘to hear’, *veure* ‘to see’) exhibit particular behavior when constructed with the infinitive form (Alsina 2002, 2423–2428; GIEC 2016, 1017–1022). In such constructions, the subject of the infinitive functions as the object (direct or indirect) of the main verb, as shown in the contrast between (13a) and (13b). In (13a) *sentir* ‘to hear’ selects a finite subordinate clause where the NP *la mare* ‘the mother’ is the subject of *plorar* ‘to cry’. Meanwhile, in (13b) the verb selects an infinitive clause and the NP *la mare* acts as the direct object of *sentir* and can be placed immediately after the main verb *sentir* or following the infinitive (13c). As a direct object, the NP may be pronominalized using an accusative clitic and agrees in gender and number with the participle of compound tenses (13d).

(13) a. He sentit que la mare plorava.
have.1.SG heard that the mother cry.IMPERF.3SG

‘I have heard *mother* crying’

b. He sentit la mare plorar.
have.1.SG heard the mother cry

‘I have heard *mother* crying’

c. He sentit plorar la mare.
have.1sg heard cry the mother

‘I have heard *mother* crying’

d. L’he sentida plorar.
her.3.SG.F.ACC have.1.SG heard.SG.F cry

‘I have heard *her* crying’

If the verb in infinitive is transitive, the subject may appear as the indirect object of the causative verb or perception verb (14a). Thus, it can be pronominalized with a dative clitic (14b). It can also act as the direct object if it occurs immediately after the

perception verb (14c), and in these cases it can be pronominalized with an accusative clitic (14d).

- (14) a. He sentit lloar el mestre a la mare.
 have.1SG heard praise the teacher to the mother
 'I've heard *mother* praising the teacher'
- b. *Li* he sentit lloar el mestre.
 3.SG.DAT have.1SG heard praise the teacher
 'I've heard *her* praising the teacher'
- c. He sentit *la mare* lloar el mestre.
 have.1.SG heard the mother praise the teacher
 'I've heard *mother* praising the teacher'
- d. *L'*he sentida lloar el mestre.
 her.3.SG.F.ACC-have.1.SG heard.SG.F praise the teacher
 'I've heard *her* praising the teacher'

h) As copular verbs, *semblar* and *parèixer* 'to seem' may occur with a subordinate clause in the subjunctive or in infinitive functioning as the subject of the main clause, and they accept a dative object: *Em sembla normal {que ens ajudi/ajudar-los}* 'It seems right to me {that (s)he helps us/to help them}'. They can also act as impersonal verbs: *Sembla que està cansat* 'It seems that he is tired'. In this case, the subordinate clause functions as the direct object and can be pronominalized with the clitic *ho* 'it': *Ho sembla* 'It seems so'. With this impersonal meaning, the dative object is also possible if the subordinate clause is in the indicative: (*Em*) *sembla que està cansat* 'It seems to me that he is tired'; however, the dative object is not possible if the subordinate clause is in the subjunctive: (**Em*) *sembla que estigui cansat* 'It seems (**to me*) that he is.SUBJ tired'. On the use of the subjunctive with *semblar*, see 4.c.

3 Relative clauses

Relative clauses are subordinate clauses introduced by a relative pronoun, a relative adverb or, more seldom, a relative determiner. Catalan has four relative pronouns: the unstressed *que*, the stressed *qui* and *què* and the compound *el qual* (article + relative pronoun). All of these are invariable, except for the compound pronoun (*el qual* [m. sg.], *la qual* [f. sg.], *els quals* [m. pl.] and *les quals* [f. pl.]), which may also function as a determiner. Moreover, there are three relative adverbs: *on* 'where', *quan* 'when' and *com* 'how'. The main properties of these relative units and relative clauses are described below (Solà 2002; GIEC 2016, 1031–1062).

a) Relative pronouns occur in clause-initial position (15a). If the pronoun is part of a phrase, then the whole phrase is also clause-initial, as with *de què* 'of/about which' (15b) where the relative pronoun *què* is the object of the preposition *de*.

- (15) a. He llegit el poema *que em recomanares*.
 'I've read the poem *that you recommended to me*'
 b. He llegit el poema *de què em vas parlar*.
 'I've read the poem *you talked to me about* (lit. about which you talked to me)'

The selection of the relative pronoun is constrained by the syntactic function that it performs within the subordinate clause, but also by the presence or absence of an antecedent, and by the type of relative clause.

b) The antecedent is the constituent of the matrix clause modified by the relative clause as a whole. In (15), the antecedent is the noun *poema* 'poem' and, thus, in (15a) the meaning of the relative clause is 'you recommended the poem to me' and in (15b) 'you talked to me about the poem'. Depending on their antecedent, relative clauses may be either restrictive or non-restrictive modifiers. In (16a), for instance, the relative clause is restrictive, since it delimits the extension of the antecedent *alumnes* 'students', whereas in (16b) it is non-restrictive, merely introducing an appositive explanation or a parenthetical aside on the antecedent *els alumnes* 'the students'.

- (16) a. Els alumnes *que no van poder fer l'examen* el faran un altre dia.
 'The students *who couldn't do the exam* will do it another day'
 b. Els alumnes, *que no van poder fer l'examen*, el faran un altre dia.
 'The students, *who couldn't do the exam*, will do it another day'

In general, the antecedent is a nominal category. It may be a noun, such as *alumnes* in (16a), a NP, such as *els alumnes* in (16b), or a pronoun, such as *algú* 'someone' in (17a) below. If the antecedent is a clause, the relative pronoun is preceded by a noun with a generic meaning like *fet* 'fact' or *cosa* 'thing' (17b). More rarely, the antecedent may be an adverb, as *allà* 'there' (17c) or a PP, as *sense mirar als ulls* 'without looking sb. in the eye' (17d).

- (17) a. Parla amb *algú que* conegui bé el tema.
 'Speak to *someone* who knows the subject well'
 b. No diuen res, *fet/cosa que* resulta preocupant.
 'They don't say a word, *which* (lit. fact/thing that) is worrying'
 c. Col·loca-ho *allà on* puguis.
 'Put it *wherever* (lit. there where) you can'
 d. Parlava *sense mirar als ulls, com* parla quan està nerviós.
 'He spoke *without looking me/us in the eye*, just as he speaks when he is nervous'

c) Because of their meaning, there are relative pronouns that can be used without an antecedent, in so-called free relative clauses. This is the case for the pronoun *qui*

‘who/whom’, referring to people, and the adverbs *on* ‘where’, *quan* ‘when’ and *com* ‘as/how’, referring to place, time and manner, respectively: *Parla amb qui vulguis* ‘Talk to whomever you wish’; *Fes-ho on/quan/com vulguis* ‘Do it wherever/when-ever/as you wish’. Halfway between the relative clauses with an antecedent and the free ones, there are some semi-free relative clauses, in which the relative pronoun is preceded by the definite article *el* ‘the’ or the distal demonstrative *aquell* ‘that/those’: *Els/Aquells qui hagin acabat ja se’n poden anar* ‘Those who have finished may already leave’.

d) The most frequent relative pronoun is the unstressed *que*, which occurs mainly in clauses with an antecedent. As shown in (15a) and (16) above, it is frequently used in restrictive as in non-restrictive relative clauses, and both with human antecedents and non-human ones. It cannot function as the object of a preposition, and thus it is essentially used as a subject (*la persona que més m’estima* ‘the person who loves me the most’) or as a direct object (*la persona que més estimo* ‘the person whom I love the most’). It may also function as a temporal adjunct when its antecedent is an adverb such as *ara* ‘now’, *demà* ‘tomorrow’ or *sempre* ‘always’: *ara que tinc vint anys* ‘now that I’m twenty’, *demà que és el meu aniversari* ‘tomorrow which is my birthday’, *sempre que puc* ‘whenever I can’. Without antecedent, it occurs only in semi-free relative clauses. It may be preceded by the individualizing article *el* (or the colloquial and non-standard neuter *lo*) if it refers to a non-animate entity (*Agafa el que necessites* ‘Take what you need’), and by the definite article or the distal demonstrative *aquell*, if it refers to human entities (*Els/Aquells que hagin acabat...* ‘Those who have finished...’). In this latter situation, however, the use of *qui* is considered more formal.

e) The stressed relative pronouns *què* and *qui* are used in clauses with an antecedent, but only as objects of unstressed prepositions (*a* ‘to’, *en* ‘in/on’, *de* ‘of/from’, *amb* ‘with’ and, to a lesser extent, *per* ‘by/for’). The first refers to inanimate entities and the second to human entities: *el tema de què vam parlar ahir* ‘the subject on which we talked yesterday’; *el noi amb qui he quedat a sopar* ‘the guy with whom I’ve arranged to have dinner’. The relative pronoun *qui*, as explained above, may also appear without an antecedent, and, in this context, it does not have to be preceded by a preposition: *Qui ho sàpiga que alci el braç* ‘Anyone who knows may raise his/her hand’ (or *El qui ho sàpiga...*, *Aquell qui ho sàpiga...* ‘The person who knows...’); *Ho pots comentar a qui consideris convenient* ‘You can comment to whomever you wish’.

f) The compound relative pronoun *el qual* may alternate with the unstressed *que* and the stressed *què* and *qui* in many contexts. Although *que* is the most usual choice, *el qual* may also occur in non-restrictive relative clauses functioning as a subject (18a) or a direct object (18b):

- (18) a. Se’n va a l’aeroport a buscar el seu fill, {que/el qual} arriba d’Austràlia.
 ‘(S)he is going to the airport to pick up his/her son, who is arriving from
 Australia

- b. He llegit de nou la seva primera novel·la, {que/la qual} em va captivar d'adolescent.
'I've reread her/his first novel, which fascinated me when I was a teenager'

Furthermore, the compound relative pronoun may appear as the object of any preposition: with unstressed prepositions, like *què* and *qui* (19a), but also with stressed ones (19b) and with prepositional locutions (19c).

- (19) a. el llibre *del qual* t'he parlat
'the book *about whom* I've talked to you'
b. les malalties *contra les quals* lluitem
'the diseases *against which* we are fighting'
c. la taula *al voltant de la qual* es van col·locar
'the table *around which* they sat'

In these contexts, the compound relative pronoun *el que* is also used colloquially instead of *el qual* (*el llibre del que t'he parlat*). However, this form is not accepted as standard because it is considered to be a calque from Spanish.

g) The compound relative pronoun *el qual* performs other particular functions not shared by the other relative pronouns. In Catalan, there is no possessive determiner equivalent to the Spanish *cuyo*, the French *dont* or the English *whose*. The relative pronoun with an equivalent function used in formal registers is *el qual* preceded by the preposition *de* and placed after the noun being modified (20).

- (20) un llibre *l'autor* *del* *qual* *ara* *no* *recordo*
a book the-author of.the which now not remember.PRES.1SG
'a book *whose author I don't remember now*'

As is customary in Catalan, in (20), the relative clause is joined to its antecedent (the noun *llibre* 'book'), but it is headed by the NP which contains the relative pronoun functioning as a modifier of this noun (*l'autor del qual* 'whose author').

h) In non-restrictive clauses, the compound relative pronoun may also function as a determiner of a noun that repeats the antecedent literally or of an anaphoric general noun: *Va adquirir unes joies, les quals joies té ara la filla* '(S)he bought some jewels, which (jewels) his/her daughter owns now'; *El van condemnar a vint anys de presó, la qual pena tothom va considerar excessiva* 'He was sentenced to twenty years in prison, a conviction which everybody considered excessive'. In this usage, the antecedent may also be a clause: *Van arribar tard, {la qual cosa/el qual fet} no va estranyar a ningú* 'They arrived late, which didn't surprise anybody'.

i) Relative adverbs usually appear without an antecedent: *He anat on em vas dir, quan he pogut i com he sabut* 'I went where you told me, when I could and as I knew how'. However, they also accept it more or less easily, depending on each relative

adverb. The least restricted is the locative on ‘where’, which can have a noun (21a), a NP (21b) or an adverb of place (21c) as its antecedent.

- (21) a. No he tornat al poble *on vaig nèixer*.
 ‘I’ve never gone back to the village *where I was born*’
 b. Arribàrem tard al peu de la muntanya, *on ens esperava l’altre grup d’esca-*
 ladors.
 ‘We arrived late at the bottom of the mountain, *where the other group of*
 climbers was waiting for us’
 c. Sí, anem allí *on ens vam conèixer*.
 ‘Yes, we’re going to the place *where we met*’

The relative adverb *on* can express location or direction, without a preposition or preceded by *a* ‘to’, although the use without *a* is preferable: *la casa (a) on visc* ‘the house where I’m living’; *la ciutat (a) on vaig* ‘the city where I’m going to’. It may also act as an object of other prepositions with spatial meaning: *el refugi des d’on hem sortit* ‘the refuge from where we have left’; *la ciutat cap on ens dirigim* ‘the city towards which we are heading’.

The modal relative adverb *com* may appear in restrictive clauses (*No m’agrada la manera com ens mira* ‘I don’t like the way (s)he looks at us’) and in non-restrictives ones (*Se n’ha anat sense fer soroll, com fa sempre* ‘(S)he left quietly, as always’). In the first case, as well as the nouns *manera* ‘manner’ or *forma* ‘form’, it may have adverbs such as *igual* or *tal* ‘in the same way’ as an antecedent: *Parla igual/tal com parlava son pare* ‘(S)he speaks in the same way as his/her father did’. The relative adverb *quan* can only have an antecedent in non-restrictive clauses: *M’ho van dir aleshores, quan ja era massa tard* ‘They told me then, when it was too late’.

j) As in other Romance languages, analytical relative clauses are frequent in colloquial Catalan (Solà 2002, 2520–2527; GIEC 2016, 1056–1057). These constructions have *que* in initial position, which can be interpreted as a mere conjunction, since the syntactic function which would correspond to the relative pronoun is performed by a clitic pronoun (22a) or a possessive determiner (22b).

- (22) a. el noi *que* li hem fet la comanda
 the guy that 3SG.DAT have.1PL make the.F order
 ‘the guy *to whom* we gave our order’
 b. la noia *que* el seu pare és metge
 the.F girl that the POSS.3SG father is doctor
 ‘the girl *whose* father is a doctor’

k) As shown in the examples analyzed up to now, relative clauses are generally speaking finite clauses. Albeit with many restrictions, relative clauses in infinitive are also possible in such cases where the relative pronoun does not act as the subject or

the direct object (Solà 2002, 2509–2511; GIEC 2016, 1061–1062). For instance, in (23a), where the relative acts as the indirect object and the implicit subject of the relative clause coincides with that of the main clause, or in (23b), where the relative is a locative adverb and the subject has a generic or indefinite interpretation.

- (23) a. No vam veure ningú a qui demanar ajut
 ‘We didn’t meet anybody to ask for help’
 b. En aquest poble no hi ha cap hostel on passar la nit
 ‘There is no hostel in this village to spend the night’

Relative clauses in infinitive express modal meanings, such as possibility, necessity, etc., which frequently may be strengthened by the auxiliary *poder* ‘can/may’: *un hostel on poder passar la nit* ‘a hostel where it is possible to spend the night’.

4 Subordination and mood

The indicative is the unmarked mood as opposed to the subjunctive. Syntactically, the indicative may appear in simple sentences and in subordinate clauses, whereas the subjunctive occurs mainly in subordinate constructions. Semantically, the indicative is featured as the mood of pure assertion, whereas the subjunctive introduces modal meanings related to subjectivity, possibility, unreality, desire, etc. The non-assertive nature of the subjunctive can be related to three types of context features, depending on the modal element which selects the subjunctive and on the resulting meaning: optative contexts, potential contexts and emotive or presuppositional contexts (Pérez Saldanya 1988; Hualde 1992, 316–323; Quer 1998; Wheeler/Yates/Dols 1999, 373–392; GIEC 2016, 932–942). However, although the subjunctive has to appear in these contexts, the indicative is not excluded, since the selection of the mood is sometimes constrained by the subordinator. In the following sections, the diverse contexts where the subjunctive is used are analyzed, contrasting them with the indicative when necessary.

a) As for the optative context, the subjunctive occurs in content clauses depending on predicates of willingness, desire (24a), command (24b), necessity (24c) or causation (24d). In general, the clause in the subjunctive refers to situations that have not yet taken place at the reference point or are presented as non-real (24a–c), but it may also introduce real situations, as in subordinate clauses depending on implicative verbs (24d).

- (24) a. Vol que *tornis* demà.
 ‘(S)he wants you *to come back* tomorrow’
 b. Ens van ordenar que no *diguéssim* res a ningú.
 ‘They commanded us not *to say* anything to anybody’

- c. És necessari que *acabeu* demà.
'It is necessary that you *finish* tomorrow'
- d. Això fa que tot *sigui* més difícil; Hem aconseguit que ens *donin* una subvenció.
'That makes it more difficult'; 'We've succeeded in getting them to *give* us a subsidy'

Verbs of saying (*dir* 'to say, to tell', *suggerir* 'to suggest', *repetir* 'to repeat', etc.) select the indicative if they assert a state of affairs or event (*Diu que ell mateix acabarà el treball* 'He says that he himself will finish the work') and the subjunctive if they express the will or command of the subject of the main verb (*Diu que acabis tu el treball* '(S)he says that you should/have to finish the work').

b) The optative context may also explain the subjunctive in other constructions such as: non-restrictive relative clauses where the speaker expresses a wish (25a), purpose clauses (25b) or illative clauses introduced by *d'aquí (ve) que* 'hence that' (25c).

- (25) a. Ens recordem molt de l'àvia, que al cel *sigui*.
'We remember our grandmother so much, may she *be* in heaven'
- b. Parla alt perquè tothom l'*entengui*
'(S)he speaks aloud so that everybody *understands* him/her'
- c. Parla alt; d'aquí que tots l'*entenguin*
'(S)he speaks aloud; hence, everybody *understands* him/her'

Independent sentences are constructed with the subjunctive when expressing wishes, preceded by the conjunction *que* or not (26a), mitigated commands or suggestions, preceded by *potser que* 'perhaps that' (26b), and prohibitions (26c).

- (26) a. Que *tinguis* bon viatge!; Tant de bo *plagui* al cel que ens retrobem aviat!
'Have a nice trip'; 'May heaven *grant* that we may meet again soon!'
- b. Potser que ho *féssiu* vosaltres.
'Perhaps you *could do* it'
- c. No *diguis* això!
'Don't *say* that!'

c) As for potential contexts, the subjunctive is used in content clauses selected by predicates expressing probability or possibility (27a), doubt (27b) or denial or a statement of falsehood (27c).

- (27) a. És possible/probable que *tingui* raó.
'It's possible/probable that you *are* right'
- b. Dubto que *hagi votat* a favor.
'I doubt that (s)he *voted* for it'

- c. És fals que tot *estigui* malament.
 'It is not true that everything *is* bad'

Verbs of saying, opinion and judgement usually select the subjunctive when negated (28a), but also accept the indicative if negation does not affect the subordinate clause (28b). The same can be applied to *negar* 'to deny' or *ignorar* 'to ignore', which include the negative meaning lexically.

- (28) a. No diu/creu que *sigueu* vosaltres els culpables.
 '(S)he is not saying/does not believe that you *are* guilty'
 b. No diu/creu que *heu* estat vosaltres els autors.
 '(S)he is not saying/does not believe that you *were* the authors'

The verb *semblar* 'to seem' may select the indicative if it means 'belief' (29a) or the subjunctive if it means 'appearance' (29b). Verbs expressing 'fear' may select both moods without any change of meaning and, in the subjunctive, accept the expletive negation (29c).

- (29) a. Sembla que *està* dolgut pels teus comentaris.
 'It seems that he *is* hurt by your comments'
 b. Sembla que *estigui* dolgut pels teus comentaris.
 'It seems that he *is* hurt by your comments'
 c. Té por que {*plourà/no plogui*} demà.
 '(S)he is afraid that *it will rain* tomorrow'

Despite being potential contexts, indirect interrogative clauses introduced by the conjunction *si* 'whether' are constructed with the indicative (30a). However, partial interrogative clauses may appear in the subjunctive if they refer to possible situations (30b–c).

- (30) a. No sé *si* enguany podrem fer vacances.
 'I don't know *whether* we will be able to go on holiday this year'
 b. Això dependrà de qui t'ho *ha/hagi* dit.
 'That will depend on who *told* you'
 c. No m'importa què *has/hagis* pogut fer.
 'I don't care what you *may have* done'

d) The potential context also allows the subjunctive in most relative clauses. In clauses with an antecedent, the subjunctive is associated to the non-specificity of the phrase in which the relative clause is embedded. In (31a), the indicative (*té* 'has') indicates that the NP *un llibre* 'a book' is specific and denotes a particular referent in the discourse universe, whereas the subjunctive (*tingui* 'has') designates a non-

specific and non-individualized referent. The same opposition is established in free relative clauses (31b). However, in prospective contexts, both the indicative, the most traditional use, and the subjunctive may be used equivalently (31c).

- (31) a. Busco un llibre que *té/tingui* il·lustracions de Joan Miró.
 ‘I’m looking for a book which has illustrations by Joan Miró’
 b. Els qui *han/hagin* acabat, ja se’n poden anar.
 ‘Those who have finished can already leave’
 c. Ho farà qui *podrà/pugui*, on *podrà/pugui*, com *podrà/pugui* i quan *podrà/pugui*.
 ‘Whoever *can* will do it, wherever he *can*, however he *can* and whenever he *can*’

e) As for adjunct constructions, conditional clauses and concessive conditional ones refer to potential or non-real situations and they should be constructed with the subjunctive. However, the mood is here constrained by the conjunction. They are constructed with the subjunctive when introduced by subordinators other than *si* ‘if’, as in the conditional clause of (32a) and the concessive conditional clause of (32b). Clauses introduced by *si* accept both moods, according to some distinctions analyzed in 8.c.

- (32) a. En cas que el pare *torni*, aviseu-me.
 ‘Just in case father *comes* back, tell me’
 b. Encara que el pare *torni* tard, aviseu-me.
 ‘Even if father *comes* back late, tell me’

Causal clauses are also constructed in the subjunctive if they are within the scope of negation. Let us remark, for instance, the contrast between (33a), where the causal clause is not affected by the negation and, thus, is constructed using the indicative and expresses an effective cause, and (33b), where the causal clause is affected by the negation, is therefore constructed with the subjunctive and expresses a non-effective cause. The subjunctive is also possible if the causal clause is within the scope of other modal elements with potential meaning, such as the predicate *és possible* ‘is possible’ (33c) or the interrogative modality (33d).

- (33) a. No va assistir a la festa perquè es *trobava* malament.
 ‘(S)he didn’t go to the party because (s)he *felt* ill’
 b. No va assistir a la festa perquè es *trobés* malament (sinó per una altra raó).
 ‘(S)he didn’t go to the party because of *feeling* ill (but for another reason)’
 c. És possible que us hagi traït perquè *estigués* dolgut.
 ‘It is possible that he betrayed you because he *felt* hurt’

- d. Van arribar tard perquè hi *hagués* molt de trànsit?
 ‘Did they arrive late because there *was* a lot of traffic?’

f) The subjunctive in sentences introduced by an adverb expressing doubt or uncertainty is recent and attributed to Spanish interference: *Potser sigui una errada* ‘Maybe it is an error’. In this case, the preferable option is to use the indicative: *Potser és una errada* ‘Maybe it is an error’.

g) The third widespread context in which the subjunctive is used is the emotive or presuppositional. The subjunctive occurs in content clauses dependent on psychological verbs that express emotions and presuppose the truth of the propositional content of the subordinate clause (34a, b). However, the indicative is also possible to highlight the informative relevance of the subordinate clause, for instance, by means of cleft sentences (34c).

- (34) a. M’alegra que *hagis obtingut* el premi.
 ‘I’m pleased that you *have won* the prize’
 b. Em sap greu que se n’*hagi* d’anar de seguida.
 ‘It’s a pity that (s)he *should have* to leave immediately’
 c. El que em sap més greu és que se n’*ha hagut* d’anar de seguida.
 ‘What I regret most is that (s)he *has had* to leave immediately’

The subjunctive may also be used in factual subordinate clauses containing given information, as in content clauses that may be preceded by the phrase *el fet* ‘the fact’ (35a) or in concessive adjuncts, as in (35b).

- (35) a. (El fet) que *hagi acceptat* parlar amb tu és molt significatiu.
 ‘(The fact) that (s)he *has agreed* to speak with you is rather significant’
 b. Ja sé que en Carles ha confessat, però encara que *hagi confessat*, ho pot negar després.
 ‘I know that Charles has confessed, but, despite *having confessed*, he can withdraw his confession afterwards’

5 Temporal constructions

Traditionally, all temporal adjuncts containing or consisting of a finite or non-finite subordinate clause are considered temporal subordinate clauses (Solà i Pujols 2002; GIEC 2016, 1189–1202). They are a heterogeneous group of constructions consisting of: relative clauses (36a); subordinate clauses introduced by a conjunction (36b); phrases with a preposition that selects as its complement a finite content clause or an infinitive one (36c,d); prepositional phrases with *en* ‘in/on’ or *al* ‘to the’ followed by an infinitive clause (36e), and gerund (36f) or participle constructions (36g).

- (36) a. Ja en parlarem *quan et trobis millor*.
 ‘We’ll speak *when you feel better*’
 b. Se’n van anar *mentre/que plovia*.
 ‘(S)he left *while it was raining*’
 c. Arribarem {*després que acabi el partit/després de sopar*}.
 ‘We’ll arrive {*after the match has finished/after having dinner*}’
 d. Deixa-ho tot endreçat {*abans que arribi la mare/abans d’anar-te’n*}.
 ‘Put in an order {*before our mother arrives home/before leaving*}’
 e. *En eixir el sol*, surten a passejar; els vaig veure *al passar per la plaça*.
 ‘*When the sun rises*, they go out for a walk; I saw them *when walking across the square*’
 f. En vam parlar (*tot/bo i*) *baixant les escales*.
 ‘We were (just) talking about it *while going down the stairs*’
 g. {*Un cop/Després de*} *presentades les esmenes*, es procedí a la votació.
 ‘{*Once submitted/After submitting*} *the amendments*, the vote was held’

Gerund and participle constructions do not require any conjunction, but some elements reinforcing their aspectual meaning may be placed in clause-initial position: gerund constructions may be preceded by the quantifier *tot* (lit. ‘all’) or the locution *bo i* (lit. ‘good and’), which intensify their meaning of temporal simultaneity in relation to the main clause (36f). As for participle constructions, they may be preceded by expressions such as *un cop* or *una vegada* ‘once’ or the preposition *després* (*de*) ‘after’, which highlight the temporal value of anteriority (36g). Conversely, finite and infinitive constructions have to be introduced by a subordinator that underlines the meaning of sequencing or simultaneity (36a, e), simultaneity and duration (36b), anteriority (36c) or posteriority (36d) of the subordinate clause in relation to the main clause.

As shown in (36e), in the most traditional usage, constructions with *en* + infinitive express sequencing (‘the sun rises and then they go out to walk’), whereas paralleled constructions with *al* + infinitive express simultaneity and duration (‘when walking across the square’). In many dialects of peninsular Catalan, however, the construction with *al* has replaced the construction with *en* (Rigau 1993; 1995; Martines 2000).

There are also impersonal temporal constructions with the verb *fer* (lit. ‘do’). They may function either as subordinate clauses (37a) or as main clauses selecting a subordinate clause with *que* or a PP with the preposition (*des*) *de* ‘since/for’ (37b, c). See Rigau (2001).

- (37) a. Es van casar *fa temps*.
 ‘They married *a long time ago*’
 b. *Fa temps* {*que es van casar/(des) del seu casament*}.
 ‘*It’s been a long time* {*since they married/since their marriage*}’

- c. *Fa temps* que és a l'hospital.
'(S)he has been in the hospital *for a long time*'

These constructions locate situations expressing the temporal distance between the reference point and the previous event point, i.e. the time when a particular event took place (37a–b) or a durative state of affairs began (37c). The temporal distance between the reference point and the event point is expressed by means of the object of the verb *fer*. This object consists of a NP with a bare or quantified temporal noun: *fa (molt de) temps* 'it's been (quite a) long time'; *feia (massa) estona* 'it's been (quite a) long while'; *farà (tres) setmanes* 'it will be (three) weeks'. The reference point may coincide with the speech point, i.e. the time of the utterance, or with a different temporal point, depending on the tense of the verb *fer* and on the temporal expression that may precede the verb. It corresponds to the speech point in (37) above or in (38a), and does not coincide with it in (38b, c).

- (38) a. *(Ara) fa* un moment que ha sortit de casa.
'(S)he left home *(just) a moment ago*'
b. *Demà farà* trenta anys que ens vam casar.
'*It will be* thirty years *tomorrow* since we married'
c. *Quan ens vam conèixer feia* temps que havia acabat els estudis.
'*When we met, it had been* some time since (s)he had finished his/her studies'

6 Clauses of degree: comparative, superlative and result constructions

Most typical comparative constructions are similar to the superlative and the result ones in that they are governed by a comparative quantifier (*més* 'more', *menys* 'less', *tan(t)* 'so/as (much/many)') and have a coda that depends on it and is introduced by a conjunction (*que* 'than/that', *com* 'as') or by a preposition (*de* 'of/in'). In comparative clauses, a relation of equality or inequality is established (39a), in superlative constructions a property in its extreme degree is attributed (39b) and in result constructions the result of an extreme situation is expressed (39c).

- (39) a. Enguany ha fet *tanta* calor *com l'any* passat.
'This year has been *as hot as last year*'
b. És l'any *més* calurós *de la darrera* dècada.
'This year is the hottest *of the last decade*'
c. Ha fet *tanta* calor *que s'han cremat* les flors del castanyer.
'It has been *so hot that the chestnut flowers have burnt*'

The governor in these constructions may also be a synthetic comparative, i.e. an adjective or adverb ending with *-or* that incorporates lexically the meaning of the quantifier *més* (*millor* ‘best, better’, *pitjor* ‘worst, worse’, *major* ‘bigger, greater, elder’, *menor* ‘smaller’): *Parla alemany* {*millor/més bé*} *que tu* ‘(S)he speaks German better than you (do)’; *És el* {*millor/més bo*} *de la classe* ‘(S)he is the best in the class’.

6.1 Comparative constructions

The relations of equality or inequality established in a comparative construction may be either quantitative or qualitative (GIEC 2016, 1071–1100; Dols/Mansell 2017, 182–188), as shown in Table 1. In the examples of Table 1, the comparison governor is marked in italics.

Table 1: Types of comparative constructions.

QUANTITATIVE			QUALITATIVE
EQUALITY		Enguany fa <i>tanta</i> calor com l’any passat. ‘This year is <i>as</i> hot as last year’	Duia un vestit <i>com</i> el teu. ‘She wore a dress <i>like</i> yours’ Duia el <i>mateix</i> vestit que tu. ‘She wore the <i>same</i> dress as you’
INEQUALITY	SUPERIOR	Enguany fa <i>més</i> calor que l’any passat. ‘This year is <i>hotter</i> than last year’	Duia un vestit <i>diferent</i> del teu. ‘She wore a <i>different</i> dress to yours’
	INFERIOR	Enguany fa <i>menys</i> calor que l’any passat. ‘This year is <i>less</i> hot than last year’	Duia un vestit <i>altre</i> que el teu. ‘She wore a <i>different</i> dress to yours’ (lit. ‘... <i>other</i> than yours’)

Comparative constructions exhibit an idiosyncratic syntactic structure and specific constituents. They are exemplified in Table 2 on the basis of the examples in Table 1 with the quantifier *més* ‘more’ and the adjective *mateix* ‘same’.

Table 2: Constituents of comparative constructions.

CONSTITUENTS	QUANTITATIVE	QUALITATIVE
governor	<i>més</i> ‘more’	<i>mateix</i> ‘same’
nucleus	<i>calor</i> ‘hot’ (lit. ‘heat’)	<i>vestit</i> ‘dress’
head	<i>més calor</i> ‘hotter’ (lit. ‘more heat’)	<i>mateix vestit</i> ‘same dress’
primary term	<i>enguany</i> ‘this year’	(<i>ella</i>) ‘she’

CONSTITUENTS	QUANTITATIVE	QUALITATIVE
secondary term	<i>l'any passat</i> 'last year'	<i>tu</i> 'you/yours'
coda	<i>que l'any passat</i> 'than last year'	<i>que tu</i> 'as you/from yours'

a) The comparative governor (*més* 'more' and *mateix* 'same') is the constituent that expresses the relation of equality or inequality. In quantitative comparatives, the governor is a degree quantifier or a synthetic comparative, whereas in qualitative ones it is an adjective or an adverb expressing identity or similarity, or the lack of identity or similarity.

b) The governor (*més* and *mateix*) modifies the constituent that acts as the nucleus (*calor*, *vestit*) and, in general, expresses the notion that is compared (degrees of heat and similarity between two dresses, respectively). The comparison is established between two terms located in different syntactic positions: the primary term (*enguany* 'this year' and the null subject *ella* 'she') is placed in a hierarchically higher position and the secondary term is integrated in the comparative coda. The coda acts as the comparative complement and is introduced by a conjunction or a preposition.

c) Comparatives expressing quantitative inequality may also contain a differential, i.e. a quantifier or quantifying locution (*molt* 'much/many', *una mica* 'a (little) bit/a few', etc.) that modifies the comparative head and highlights the difference established between the compared terms with respect to the compared notion: *molta més calor* 'much hotter', *una mica més de calor* 'a little bit hotter'.

d) The only obligatory element in comparative constructions is the governor, since the rest may remain implicit. In *Enguany menja molta més carn que l'any passat* '(S)he eats much more meat this year than last year', all the constituents occur, but in *Enguany menja més carn* '(S)he eats more meat this year', the differential and the coda is implicit; in *Menja més carn* '(S)he eats more meat', the primary term does not occur either, and in *Menja més* '(S)he eats more', neither does the nucleus. It is also possible for a constituent to assume different functions, as in *Menja més carn que peix* '(S)he eats more meat than fish', where *carn* 'meat' is both the nucleus and the primary term. Moreover, the terms of comparison may also be multiple, as in *Ella menja més carn que tu peix* 'She eats more meat than you (eat) fish', where *ella* 'she' and *carn* 'meat' function as the primary term and *tu* 'you.SG' and *peix* 'fish' as the second.

e) Comparatives of inferiority are less common than those expressing superiority. Instead of a comparative of inferiority, a negative comparative of equality or a comparative of superiority with *més poc* 'much/many less' is sometimes used: *Enguany no ha plogut tant com l'any passat* 'It hasn't rained as much this year as last year', *Enguany ha plogut més poc que l'any passat* 'It has rained much less this year than last year'.

f) In quantitative comparatives, the coda may be phrasal or clausal depending on various factors (Hualde 1992, 208–214; Wheeler/Yates/Dols 1999, 95–102). In Table 3

these factors are summarised and the types of coda are exemplified on the basis of a comparative of equality and one of inequality. The terms of comparison in the examples are in italics and the nucleus in small capitals.

Table 3: Types of comparative codas.

CODA	CONSTRUCTION PROPERTIES	EXAMPLE
PHRASAL	the terms of comparison are simple the primary term is different from the nucleus	<i>L'Aina és tan ALTA com tu.</i> 'Aina is as TALL as you' <i>L'Aina és més ALTA que tu.</i> 'Aina is TALLer than you'
	the terms of comparison are simple the primary term is the nucleus	<i>Menja tanta CARN com peix.</i> 'He eats as much MEAT as fish' <i>Menja més CARN que peix.</i> 'He eats more MEAT than fish'
CLAUSAL	the terms of comparison are multiple	<i>Ella menja tanta CARN com tu (menges) peix.</i> 'She eats as much MEAT as you (eat) fish' <i>Ella menja més CARN que tu (menges) peix.</i> 'She eats more MEAT than you (eat) fish'
	the terms of comparison are simple the primary term is the nucleus the compared element is the same in both items the secondary term is not explicit in the coda	<i>Compra tants LLIBRES com pot llegir.</i> 'He buys as many BOOKS as he can read' <i>Compra més LLIBRES {dels que necessita/que no necessita}.</i> 'He buys more BOOKS than he needs'

In comparatives of equality, the coda always occurs with the conjunction *com* 'as'. In those of inequality it occurs with *que* 'than', with the exception of the last example in Table 3, where both *que* 'than' and *de* (lit. 'of') are possible.

In codas with *de*, this preposition is followed by a NP with a relative clause where the antecedent is implicit and may be retrieved from the primary term ('of the books that he needs'). They may also be constructed with certain modal adjectives (*més llibres dels necessaris/convenients* 'more books than necessary/suitable', lit. 'of the necessary/suitable') or with the NP *el compte* 'the count' (*més llibres del compte* 'more books than necessary'). All the codas with *que* accept or require the expletive negation, with *no* or *no pas*, in contrast to the codas with *de*, which do not accept it: *més que (no (pas)) tu* 'more than you' (lit. 'more than not you'); *més carn que (no (pas)) peix* 'more meat than fish' (lit. 'more meat than not fish'), *més que no (pas) necessita* 'more than he needs' (lit. 'more than he not needs').

g) Correlative comparative constructions express the parallel increasing or decreasing of two magnitudes. These constructions consist of a complex sentence where the main and subordinate clauses contain a quantifier of inequality (or a synthetic

comparative) before the verb; the subordinate clause precedes the main clause and is introduced by *com* ‘as’:

- (40)a. *Com més difícil és el repte, més ganes té d’implicar-s’hi.*
 ‘The more difficult the challenge is, the more (s)he wants to be involved in it’
 b. *Com menys coses sàpigues, menys patiràs*
 ‘The less you know, the less you will suffer’
 c. *Com més li ho repeteixis, pitjor (serà).*
 ‘The more you repeat it to him/her, the worse it will be’

h) Among qualitative comparatives, those constructed with the relative adverb *com* have to be distinguished from the rest. In the former, the relative adverb underlines the relation of equality and introduces the coda. The other constructions exhibit an adjective or an adverb selecting the comparative complement and the type of subordinator, as shown in Table 4.

Table 4: Governors and subordinators in qualitative comparatives.

COMPARATIVE	GOVERNOR AND SUBORDINATOR	EXAMPLE
Equality	<i>com</i> ‘as/like’	una proposta <i>com</i> la d’ahir ‘a proposal like that from yesterday’
	<i>igual... que/com/a</i> ‘the same... as’	una proposta <i>igual que/com/a</i> la d’ahir ‘the same proposal as that from yesterday’
	<i>mateix... que/de</i> ‘the same... as’	la mateixa proposta { <i>que</i> ahir/d’ahir} ‘the same proposal as yesterday’
	<i>semblant/paregut a</i> ‘similar to’ <i>idèntic a</i> ‘identical/equal to’	una proposta <i>semblant a</i> la d’ahir ‘a proposal similar to that from yesterday’
Inequality	<i>diferent a/que</i> <i>distint a/que</i> ‘different/distinct from’	una proposta <i>diferent {a</i> la d’ahir/ <i>que</i> la d’ahir} ‘a proposal different from that from yesterday’
	<i>altre que</i> (lit. ‘other than’)	una proposta <i>altre que</i> la d’ahir ‘a proposal different from yesterday’s’ (lit. ‘a proposal other than the of yesterday’)

With the exception of *com* ‘as/like’ and *altre* ‘other’, qualitative governors designate a symmetrical feature and the compared elements may be expressed by means of both comparative items (*la proposta d’avui és semblant a la d’ahir* ‘Today’s proposal is similar to that from yesterday’), but also by means of coordinate phrases (*la proposta d’ahir i la d’avui són semblants* ‘Yesterday’s and today’s proposals are

similar’) or with a plural phrase (*les propostes són semblants* ‘the proposals are similar’).

6.2 Superlative constructions

Superlative constructions express comparison between the members of a whole set. In these constructions a member of the whole set is presented as having a property in its extreme degree in relation to the other members. For instance, in (41), the property of being the cleverest student in relation to the class that she belongs to (41a), or in relation to all the students that the speaker knows (41b).

- (41) a. L’alumna més intel·ligent de la seva classe.
 ‘The cleverest student in her class’
 b. L’alumna més intel·ligent que conec.
 ‘The cleverest student that I know’

Superlative constructions generally occur in definite NPs which contain a comparative quantifier of inequality (*més* ‘more’ in the previous examples) and a coda introduced by the preposition *de* or the conjunction *que*: *de la seva classe* ‘in her class’ in (41a) and *que conec* ‘that I know’ in (41b). The primary term is identified with the nucleus of the NP. In general, this is a noun (*alumna* ‘student.F’ in (41)), but may also be a relative or interrogative pronoun (*qui* ‘who’ in *És qui més ens ha ajudat de tots* ‘He is the one who has helped us most of all’) or the individualizer article *el* (*el més interessant de tot* ‘the most interesting of all’). On the other hand, the whole set for comparison is identified by the NP contained in the coda: *la seva classe* ‘her class’ in (41a) and (*les alumnes*) *que conec* ‘(the students) that I know’ in (41b).

The quantified phrase may modify the primary term directly (42a) or may be integrated within a participle construction (42b), a PP (42c) or a relative clause (42d).

- (42) a. el llibre [*més interessant*] d’aquest autor.
 ‘The *most interesting* book by this author’
 b. el llibre [*més ben editat*] d’aquest autor.
 ‘The most *well edited* book by this author’
 c. el llibre [*amb més il·lustracions*] d’aquest autor.
 ‘The book by this author with *the most illustrations*’
 d. el llibre [*que més m’ha agradat*] d’aquest autor.
 ‘The book by this author that *I’ve most enjoyed*’

6.3 Result constructions

The result constructions analyzed in this section are degree constructions and express the result of an extreme situation. See section 7.3 for other result constructions. Thus, in *Últimament la Neus fa tant d'exercici que s'ha aprimat molt* 'Lately, Neus has done so much exercise that she has lost a lot of weight' the situation of Neus doing exercise is presented to an extreme degree ('she does a lot of exercise') and this situation provokes the result that she has lost a lot of weight. In result constructions, the coda is clausal and introduced by the conjunction *que*: *que s'ha aprimat molt* 'that she has lost a lot of weight'. The extreme degree is expressed by means of the quantifier of equality *tan(t)* 'so much/many' if it refers to the superior degree as in the previous example, or by the same quantifier followed by *poc* 'little/few' if it refers to the inferior degree: *La Neus fa tan poc d'exercici que s'ha engreixat molt* 'Neus does so little exercise that she has gained a lot of weight'.

The extreme degree may also be expressed by other elements: the demonstrative *tal* 'such' (43a), the universal quantifier *cada* 'each' (43b) or an exclamatory or interrogative quantifier (43c).

- (43) a. Ens va contestar de *tal* manera que ens en vam anar.
 '(S)he answered in *such* a manner that we left'
 b. Té *cada* reacció que fa por.
 '(S)he has *such* reactions that (s)he frightens everybody'
 c. *Quant* que deu haver begut que no s'aguanta dret!
 '*How* much should (s)he have drunk to stumble!'

Midway between causal and result clauses, there are certain adjuncts introduced by the prepositions *amb* 'with' or *de* 'of'. As shown in (44), the adjunct expresses the situation in an extreme degree presented as the cause of what is said in the sentence to which it is attached.

- (44) a. No parava d'abraçar-nos, *de tan content que/com estava*.
 '(S)he embraced us now and again, *being so happy as (s)he was*'
 b. *Amb tant de soroll que hi havia*, no el vam sentir arribar.
 '*With so much noise*, we didn't hear him/her arriving'

In these adjuncts, the quantified phrase with *tan(t)* occurs focalized between the prepositions *de* or *amb* and the conjunction *que* (or more unusually *com*) that introduces the rest of the clause if the verb is in a finite form.

7 Causal, purpose and illative constructions

Causal, purpose and consequence or illative constructions express cause-effect relations: cause in the case of causal clauses, purpose or destination in purpose ones and effect in illative structures (Viana/Suñils 2002, 2939–2952; Pérez Saldanya 2015; GIEC 2016, 1110–1122).

7.1 Causal constructions

Most typical causal constructions express the cause or the reason for a given situation (45a). They can also introduce an explanation, as in (45b), in which the causal clause justifies the speaker's assertion expressed in the main clause.

- (45) a. Els camps estan tots negats *perquè ha plogut molt*.
 ‘All the fields are flooded *because it has been raining a lot*’
 b. Ha plogut molt, *perquè els camps estan tots negats*.
 ‘It has been raining a lot, *since/as/for all the fields are flooded*’

Formally, they may be finite clauses introduced by a conjunction, as *perquè* ‘because’ in the previous examples in (45), or by a causal preposition, as *per* ‘lit. for’ or *a causa de* ‘because of’ in (46a, b). In the latter case, the complement of the preposition may be a NP (46a) or, with more restrictions, an infinitive clause (46b). Causal clauses may also be absolute participle constructions, more or less lexicalized, with a NP (46c) or a content clause with the conjunction *que* (46d) as grammatical subjects.

- (46) a. S’ha anul·lat el concert {*per/a causa de*} *la tempesta*.
 ‘The concert has been cancelled *because of the storm*’
 b. Li han retirat el carnet *per* {*conduir/haver conduït*} *ebri*.
 ‘Her/his license has been suspended *because of driving/having been driving drunk*’
 c. *Vistes/Ateses les reclamacions presentades*, es va retirar la proposta.
 ‘*In view of/Regarding the submitted complaints*, the proposal was withdrawn’
 d. *Vist/Atès que s’han presentat moltes reclamacions*, s’ha retirat la proposta.
 ‘*[In view/Considering] that a lot of complaints have been submitted*, the proposal has been withdrawn’

Causal constructions may function either as internal or external adjuncts to the predicate. Internal adjuncts express motives or reasons and, when containing a finite clause, are usually constructed with the conjunction *perquè* ‘because’ (47a). Syntacti-

cally, they may be used as the answer to a question with *per què* ‘why’ (47b) and may fall within the scope of negation (47c).

- (47) a. Se’n va anar perquè es trobava malament.
 ‘(S)he left because (s)he felt ill’
 b. – Per què se’n va anar? – Perquè es trobava malament.
 ‘– Why did she leave? – Because (s)he felt ill’
 c. No se’n va anar perquè es trobés malament (sinó per una altra raó).
 ‘(S)he didn’t leave because of feeling ill (but for another reason)’

The external causals constitute sentential adjuncts and are appositive, thus present comma intonation, as in (48). Semantically, external causals are explicative. They may justify the propositional content of the whole sentence, as in (48a). But they can also have a pragmatic meaning. In this case, they may be either epistemic causals or speech act causals. The epistemic ones highlight the speaker’s assumption asserting the state of affairs expressed in the main clause, as in (48b), where the causal clause justifies why the speaker asserts that Peter has blamed us. On the other hand, speech act causal clauses highlight the illocutionary force of the main clause, as in (48c), where the causal clause accounts for the command expressed by the imperative, or in (48d), where it offers the reason for the addressor’s question.

- (48) a. Em vaig alegrar de parlar amb ella, *perquè feia molt de temps que no la veia*.
 ‘I was glad to talk to her, *for I hadn’t seen her for a long time*’
 b. Ens ha acusat en Pere, *perquè no ens mira als ulls*.
 ‘It is Peter who blamed us, *for he won’t look us in the eye*’
 c. Ves-te’n, {*perquè/que si no*} *faràs tard*.
 ‘Go away, {*because/or else*} *you will be late*’
 d. Li han donat el premi, *que està tan contenta*?
 ‘Has she got the prize? *For she is so happy*’

Moreover, depending on the sentence information structure, it is possible to establish a distinction between rhematic causals (providing new and relevant information) and thematic or presupposed causals (which provide given or secondary information). All the causal clauses exemplified in (48) above are rhematic and occur in the right periphery of the sentence. Conversely, the causals in (49) are thematic or presupposed. Thus, they usually occur in the left periphery of the sentence (49a,b), although they may also appear in the right periphery (49c). Among thematic causals, the distinction between propositional content causals (49a) and speech act causals (49b, c) may also be established.

- (49)a. *Com que feia temps que no la veia*, em vaig alegrar molt de trobar-la.
 ‘Since/As I hadn’t seen her for a long time, I was very glad to see her’
- b. *Ja que no tens res més a fer*, ajuda’m!
 ‘Since you haven’t got anything else to do, help me!’
- c. *Ajuda’m, ja que no tens res més a fer*.
 ‘Help me, since you don’t have anything else to do!’

Table 5 shows the most common subordinators for each type of causal clause and the analyzed examples corresponding to each type. Conjunctions that are obsolete but are still used in certain literary contexts are marked with “+”.

Table 5: External causal constructions.

<i>causal</i>	<i>informative value</i>	<i>Subordinator</i>	<i>Examples</i>
content	rhematic	perquè, puix (que), ja que, +car	(48a)
	thematic	com que, vist/atès que	(49a)
epistemic	rhematic	perquè, +car	(48b)
speech act	rhematic	que, perquè, +car	(48c,d)
	thematic or presupposed	ja que, vist/atès que, +puix (que)	(49b,c)

The conjunction *perquè* is the most ductile and polyfunctional, since it is used in the internal causal clauses and in most of the external ones. The other connectives are subject to more restrictions and are only used in certain types of causal constructions. Those of temporal origin (*puix (que)*, *ja que* ‘since/as’ may occur both in the thematic speech act clauses (or proximal) and in the rhematic content causals. The former meaning is the seminal one and has fallen into disuse in the case of *puix (que)* (Pérez Saldanya/Hualde 2017). The conjunction *com que* (*com* in Old Catalan) typically introduces thematic content-causal clauses, which can only occur in the left sentence periphery, in contrast to the other thematic causal constructions. The conjunction *que* is only used in the rhematic speech act causal clauses (or proximal). In this type of clause, the use of *perquè* is enhanced by the sequence *si no* ‘if not’ (*perquè/que si no...*).

7.2 Purpose constructions and other related constructions

So-called final subordinate constructions include purpose constructions, which express the purpose for which an action is performed (50a), and other constructions closely related to them which express some type of meaning related to the idea of destination: use or profit (50b), lack, sufficiency or excess (50c), necessity or conve-

nience (50d), etc. (GIEC 2016, 1123–1132). In general, they are constructed with the conjunction *perquè* ‘so that/in order that’ in finite clauses with verbs in the subjunctive, or with the prepositions *per* or *per a* ‘to/for’ followed by an infinitive clause or a NP.¹

- (50) a. Li trucaré {perquè m’expliqui/per (a) explicar-li} què ha passat.
 ‘I’ll call her/him {so that (s)he can explain/to explain to her/him} what has happened’
- b. Aquest llibre et servirà {perquè et preparis/per (a) preparar-te/per a la preparació de} les oposicions.
 ‘This book will help you {so that you can prepare/to prepare/in preparation for} the civil service examination’
- c. Falta mitja hora {perquè arribin/per (a) arribar} allà.
 ‘There is half an hour for them to arrive there/for us to arrive there’
- d. {Perquè li donin/Per (a) demanar} el préstec, necessita avals.
 ‘In order for him/her to be granted the loan/To be granted the loan}, (s)he needs endorsements’

Purpose clauses are usually adjuncts to the predicate (50a), but they may also be objects of a NP (*l’anada a Banyoles per (a) visitar la família* ‘The trip to Banyoles to visit the family’) or sentential adjuncts, in the case of speech act purpose clauses (*Per (a) ser-te sincer, no en sabia res* ‘To be honest, I didn’t know anything’). Clauses expressing destination, in their turn, may function as complements selected by certain verbs, such as *servir* ‘help/be useful’ (50b), as adjuncts to the predicate (50c) and, more unusually, as sentential adjuncts (50d). They may also be subject complements (*Això és per (a) preparar el sopar* ‘This is to cook the dinner’), complements of a NP (*utensilis per (a) cuinar* ‘utensils for cooking’) or of an adjective (*útil per (a) cuinar* ‘useful for cooking’) or correlate with a degree quantifier (*prou per (a) aprovar l’examen; massa gent per (a) preparar el sopar* ‘enough to pass the exam; too many people to cook the dinner’).

Purpose clauses may also be constructed with the prepositional locutions *per tal (de)*, *a fi (de)* ‘in order to/with the purpose of’ (51a). The use of the conjunction *que* with finite clauses and that of the preposition *a* with infinitive clauses are more restricted. The former occurs in purpose clauses located in the right periphery and expressing ‘command’, ‘advice’ or ‘permission’ (51b). The latter is used as a comple-

¹ From a diatopic perspective, *per* is used in most of the Catalan territory, though *per a* is also used in Valencian and some northwestern dialects (see 75.2 The Simple Sentence). In the standard language, it was advised to use *per* in purpose clauses with an infinitive and *per a* in those expressing destination, though both prepositions are considered admissible nowadays in all of these contexts (Fabra ⁷1933, ¹1918, § 128; GIEC 2016, 1124–1126).

ment (51c) or an adjunct to the predicate (51d) with verbs expressing motion, position change or interruption of a motion or an activity.

- (51) a. {*Per tal/A fi*} *d'evitar problemes*, convé que feu ja la reserva.
 '*In order to avoid problems*, it is advisable to book early'
- b. Tanqueu la porta, *que no entri el fred*.
 '*Close the door, so as to keep out the cold*'
- c. Hem vingut *a veure com et trobaves*.
 '*We've come to see how you feel*'
- d. {*Es va asseure/S'hi va quedar*} *a llegir el diari*.
 {(s)he sat/(s)he stayed there} *to read the newspaper*'

7.3 Illative constructions

Illative constructions express the result of what has been said in the preceding clause. They consist of conjunctive locutions followed by a finite clause in the indicative, with the exception of *d'aquí (ve) que* 'hence that', which is constructed with the subjunctive. The locutions *així que* 'so/therefore/thus' (52a) and *d'aquí (ve) que* (52b) contain a demonstrative adverb (*així* 'thus', *aquí* 'here') which maintains a certain anaphoric value, retrieving ideas expressed in the previous clause. The locutions *o sigui que* (lit. 'or be that') (52c) and *de manera/forma que* (lit. 'in such a manner/way that') (52d) also express result.

- (52) a. Feia molt de fred; *així que decidírem tornar a casa*.
 '*It was very cold, so we decided to go back home*'
- b. La proposta estava consensuada; *d'aquí que s'aprovés fàcilment*.
 '*The proposal was agreed; hence that it was approved easily*'
- c. No ens han convidat, *o sigui que no hem de patir*.
 '*We've not been invited, so we don't have to worry about it*'
- d. Vam arribar d'hora, *de manera que vam aparcar sense problemes*.
 '*We arrived early, so we parked without any problem*'

Illative constructions have to be placed after the part of the sentence expressing the cause and constitute an independent whole, which in written discourse is marked by a comma or other punctuation marks.

8 Conditional and concessive constructions

Conditional and concessive constructions also express cause-effect relations, but in this case it is a conditioned-cause relation, in the case of conditional sentences (53a), and a non-effective cause, in the case of concessive ones (53b).

- (53) a. *Si demà fa bon temps*, anirem d'excursió a la muntanya.
 'If the weather is good tomorrow, we'll go hiking in the mountains'
 b. *Encara que demà faci mal oratge*, anirem d'excursió a la muntanya.
 'Even if the weather is bad tomorrow, we'll go hiking in the mountains'

Both constructions are typically sentential adjuncts that occur in the left periphery, as in (53). However, they may also occur in the right periphery and even be inserted within the main clause predication if they afford information presented as relevant or restricting what is expressed in the predication to which they are adjoined, as in (54).

- (54) a. Només anirem d'excursió a la muntanya *si fa bon temps*.
 'We'll go hiking in the mountain *only if the weather is good*'
 b. Anirem d'excursió a la muntanya *encara que faci mal oratge*.
 'We'll go hiking in the mountain *even if the weather is bad*'

a) The conjunction *si* 'if' is the most typical in conditional constructions. It may correlate with anaphoric adverbs such as *llavors* or *aleshores* 'then', which reinforce the conditioned-effect value of the main clause (55a). The hypothetical value of *si* may be strengthened by means of PPs with the noun *cas* 'case', such as *si de cas*, *si per cas*, *si a un cas* 'just in case' (55b). It may also be preceded by a focusing adverb such as *fins i tot*, *inclús* 'even', which lends the construction a concessive conditional meaning (55c).

- (55) a. *Si em promets que t'ocuparàs tu del gosset*, *aleshores* te'l pots quedar.
 'If you promise to take care of the little dog, *then* you can keep it'
 b. *Si de cas t'ho repenses*, no dubtis a avisar-me.
 'Just in case you change your mind, don't hesitate to tell me'
 c. No hi accediré, *fins i tot si m'ho demana el president*.
 'I won't accept it, *even if* the president asks for it'

b) Besides the conjunction *si*, there are other subordinators that may express 'condition'; for instance, the preposition *amb* 'with' (56a) or the prepositional locution *en cas de/que* 'in case of/that' (56b), which accept a NP or a finite clause with the conjunction *que* as a complement. There are also adverbs of time that may be used as conditional conjunctions, either alone, as in the case of *mentre* 'while' (56c), or followed by *que*, as *sempre que* 'provided that' (lit. 'always that') (56d).

- (56) a. *Amb {una persona/que vingui una persona}, és suficient per a acabar el treball.*
 ‘With {one person/only one person coming}, it is enough to finish the work’
 b. *En cas {d’incendi/que es produeixi un incendi}, no utilitzeu l’ascensor.*
 ‘In case {of fire/there is a fire}, don’t use the lift’
 c. *Mentre no hagi de parlar amb ell, tot anirà bé.*
 ‘As long as you don’t have to speak with him, everything will be right’
 d. *Assistiré demà a l’acte, sempre que tu m’hi acompanyis.*
 ‘I’ll attend the event tomorrow, provided that you come with me’

The locution *sempre que* is used to introduce conditions presented as strong requirements, with a meaning equivalent to that expressed by the formulaic *si i només si* ‘if and only if’. The same may be applied to the locutions *a condició (de)* ‘on the condition of/that’ (57a) and, with an excluding meaning, also to the exceptive particles (*tret, llevat, excepte, fora* ‘except if/unless’) followed by a content clause with *que* (57b). Conversely, focusing adverbs such as *sols* and *només* before a subordinate with *que* express minimal requirements (57c).

- (57) a. *Signarem el contracte a condició que s’inclogui una nova clàusula.*
 ‘We’ll sign the contract on the condition that a new clause is included’
 b. *No signarem el contracte, tret que s’inclogui una nova clàusula.*
 ‘We won’t sign the contract unless a new clause is included’
 c. *Només que s’inclogui una nova clàusula, signarem el contracte.*
 ‘If only a new clause is included, we’ll sign the contract’

c) Complex sentences containing a conditional construction follow typical temporal sequences, which depend on both the temporal localization and the modality of the expressed situations (Salvador 2002, 2991–2999; GIEC 2016, 1140–1144). Considering both factors, and simplifying the possible distinctions, three broad types of conditional clauses are usually distinguished in the grammatical tradition: open conditionals, which express situations possible in the sphere of present and future; remote conditionals, which express situations presented as false or possibly false in the same sphere, and past counterfactual conditionals, which express situations that could have taken place in the past, but did not ultimately come to pass. In Table 6, the most common temporal and modal schemas for these three types of conditionals are summarized and exemplified on the basis of clauses with the conjunction *si* and with the prepositional locution *en cas (de)*.

Table 6: Sequence of tenses in conditionals.

CLASS OF CONDITIONAL	SEQUENCE OF TENSES	
	WITH <i>SI</i>	WHITH <i>EN CAS (DE)</i>
open	<i>present indicative – future</i> Si m'ho demana, li ho diré. 'If (s)he asks for it, I'll tell him/her'	<i>present subjunctive – future</i> En cas que m'ho demani, li ho diré. 'In the case that (s)he asks for it, I'll tell him/her'
remote	<i>imperfect subjunctive – conditional</i> Si m'ho demanés, li ho diria. 'If (s)he asked for it, I'd tell him/her'	<i>imperfect subjunctive – conditional</i> En cas que m'ho demanés, li ho diria. 'In the case that (s)he asked for it, I'd tell him/her'
past counterfactual	<i>pluperfect subjunctive – conditional perfect</i> Si m'ho hagués demanat, li ho hauria dit. 'If (s)he had asked for it, I'd have told him/her'	<i>pluperfect subjunctive – conditional perfect</i> Si m'ho hagués demanat, li ho hauria dit. 'In the case that (s)he had asked for it, I'd have told him/her'

In open conditionals, the subordinate clause is constructed with the indicative using the conjunction *si*, but with the subjunctive if the locution *en cas (de)* or any other conjunction different from *si* is used. Conversely, in the remote and the past counterfactual conditional, the subjunctive is used in both cases. In the former, the imperfect indicative is also possible with the conjunction *si* (*si m'ho demanava*), although nowadays it is characteristic of a rather formal or literary style.

d) Among concessive constructions, it is common to distinguish between pure concessive (or simply concessive) constructions, and concessive conditionals, which express hypothetical or non-real situations (Quer 1998, 225–261; Salvador 2002, 3009–3022; GIEC 2016, 1153–1165; Dols/Mansell 2017, 152–153). There are some subordinators that may be used in both types of constructions and some others that are linked to just one of these types. The conjunction *encara que* 'although/though' is the most prototypical and may be used in both constructions. As a concessive conjunction, it is constructed with the indicative (58a) or with a factual subjunctive (58b); as a concessive conditional conjunction, it is used in the subjunctive (58c).

- (58) a. *Encara que era molt tard*, ens van convidar a sopar.
'Although it was late, they invited us to have dinner'
 b. *Encara que fos molt tard*, ens van convidar a sopar.
'Although it was late, they invited us to have dinner'
 c. *Encara que pogués*, no ens ajudaria.
'Although (s)he could, (s)he wouldn't help us'

The preposition *malgrat* ‘in spite of’ and the prepositional locutions *a pesar*, *a desgrat*, *a despit* (*de*) ‘despite/in spite of’ have a concessive meaning. They may select as a complement a NP (59a), an infinitive clause (59b) or a finite clause introduced by *que* (59c).

- (59) a. {*Malgrat/A pesar de*} la situació en què es troba, ens va voler ajudar.
 ‘{Despite/In spite of} the situation (s)he is in, (s)he intended to help us’
 b. {*Malgrat/A pesar de*} no haver estudiat, va aprovar l’examen.
 ‘{Despite/In spite of} not having studied, (s)he passed the exam’
 c. {*Malgrat/A pesar*} que no havia estudiat, va aprovar l’examen.
 ‘Although (s)he had not studied, (s)he passed the exam’

The conjunctive locutions *si bé* (lit. ‘if well’) or (*per*) *bé que* (lit. ‘(for) well that’) (60a) also have a concessive meaning and *ni que* (lit. ‘neither that’) (60b) and *mal que* (lit. ‘bad that’) (60c) exhibit a concessive conditional meaning, though their use is restricted and colloquial.

- (60) a. {*Si bé/Bé que*} tothom hi estava d’acord, no es va aprovar la proposta.
 ‘Although everybody agreed with it, the proposal was not approved’
 b. No el perdonaria, *ni que m’ho demanés a genollons*.
 ‘I wouldn’t forgive him/her, even if (s)he begged for it on knees’
 c. *Mal que ens pesi*, hem d’actuar com si no hagués passat res.
 ‘It’s a pitiful that we have to act as if nothing has happened’

e) There are three particular types of constructions that also express concessive conditional meanings: polar, parametric and alternative concessive conditionals (Salvador 2002, 3010–3012; Quer 1998, 235–253). Polar concessive conditionals express an extreme degree meaning. They contain a focalizing quantifier, such as *molt* ‘much/many’ in (61a), or a quantified phrase, such as *més coses* ‘more things’ in (61b), which express an extreme quantity on a scale. They are followed by the conjunction *que*:

- (61) a. *Per molt que hi insisteixes*, els teus amics no et faran cas.
 ‘However much you insist on it, your friends will ignore you’
 b. *Per més coses que li diguis*, no canviarà de parer.
 ‘Whatever you tell him/her, (s)he won’t change her/his mind’

Parametric concessive conditionals consist of a construction denoting a universal meaning. They are introduced by a verb in the subjunctive, followed by a non-specific free relative (62a) or a NP with a non-specific relative clause (62b), where the verb is repeated.

- (62) a. *Parlis amb qui parlis*, escolta bé el que et diguin.
 ‘Whoever you talk to, pay attention to what (s)he tells you’
 b. *Faci el sopar que faci*, sempre t’agrada.
 ‘Regardless of the dinner that I cook, you always like it’

Finally, alternative concessive conditionals consist of the disjunction of a conditional and its negation. Specifically, there are two predicates in the subjunctive expressing opposed situations:

- (63) *Faci fred o (faci) calor*, sempre vas en màniga de camisa.
 ‘Whether it is cold or hot, you are always in your shirtsleeves’

9 Bibliography

- Alsina, Àlex (2002), *L’infinitiu*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 3, Barcelona, Empúries, 2389–2454.
- Bonet, Sebastià (2002), *Les subordinades substantives*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 3, Barcelona, Empúries, 2321–2387.
- Bonet, Sebastià/Solà, Joan (1986), *Sintaxi generativa catalana*, Barcelona, Enciclopèdia catalana.
- Cuenca, Maria Josep (1991), *L’oració composta*, vol. 2: *la subordinació*, València, Universitat de València.
- Dols, Nicolau/Mansell, Richard (2017), *Catalan: An Essential Grammar*, London/New York, Routledge.
- Fabra, Pompeu (1933, 1918), *Gramàtica catalana*, Barcelona, Institut d’Estudis Catalans.
- GIEC = Institut d’Estudis Catalans (2016), *Gramàtica de la llengua catalana*, Barcelona, Institut d’Estudis Catalans.
- Hualde, José Ignacio (1992), *Catalan*, London/New York, Routledge.
- Martines, Josep (2000), *Sobre una altra construcció catalana força controvertida (1), “en” + infinitiu temporal al País Valencià*, in: Jordi Ginebra/Raül-David Martínez Gili/Miquel Àngel Pradilla (edd.), *La lingüística de Pompeu Fabra*, vol. 2, Alacant, Institut Interuniversitari de Filologia Valenciana, 127–164.
- Pérez Saldanya, Manuel (1988), *Els sistemes modals d’indicatiu i de subjuntiu*, Barcelona, Institut de Filologia Valenciana/Publicacions de l’Abadia de Montserrat.
- Pérez Saldanya, Manuel (2015), *Les construccions causals en català: classes i nexes que les introdueixen*, *Els Marges* 105, 10–38.
- Pérez Saldanya, Manuel/Hualde, José Ignacio (2017), *From Theme to Rheme: the Evolution of Causal Conjunctions of Temporal Origin in Catalan*, *Studies in Hispanic and Lusophone Linguistics* 10/2, 319–348.
- Quer, Josep (1998), *Mood at the Interface*, The Hague, Academic Graphics.
- Rigau, Gemma (1993), *La legitimació de les construccions temporals d’infinitiu*, in: Amadeu Viana (ed.), *Sintaxi. Teoria i perspectives*, Lleida, Pagès, 231–252.
- Rigau, Gemma (1995), *The Properties of the Temporal Infinitive Constructions in Catalan and Spanish*, *Probus* 7, 279–301.
- Rigau, Gemma (2001), *Temporal Existential Constructions in Romance Languages*, in: Yves d’Hulst/Johan Rooryck/Jan Schroten (edd.), *Romance Languages and Linguistic Theory 1999*, Amsterdam/Philadelphia, Benjamins, 307–333.

- Salvador, Vicent (2002), *Les construccions condicionals i les concessives*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 3, Barcelona, Empúries, 2977–3025.
- Serra, Enric/Prunyonosa, Manuel (2002), *La coordinació*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 3, Barcelona, Empúries, 2181–2245.
- Solà, Joan (2002), *Les subordinades de relatiu*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 3, Barcelona, Empúries, 2455–2565.
- Solà i Pujols, Jaume (2002), *Modificadors temporals i aspectuals*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 3, Barcelona, Empúries, 2867–2936.
- Viana, Amadeu/Suïls, Jordi (2002), *Les construccions causals i les finals*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 3, Barcelona, Empúries, 2937–2975.
- Wheeler, Max W./Yates, Alan/Dols, Nicolau (1999), *Catalan*, London/New York, Routledge.

Ingo Feldhausen and Xavier Villalba

5.4 Modality and Information Structure: Focus, Dislocation, Interrogative and Exclamatory Sentences

Abstract: This chapter provides an overview of different sentence types and structures in Catalan that are closely related to modality and information structure in this language. In doing so, we mainly concentrate on syntactic, prosodic, and interpretative aspects of the presented structures. We first explore different concepts of focus and discuss how focus is syntactically and prosodically marked. We present a brief chronological overview of previous work and highlight dialectal differences. Dislocations and their information structural function are introduced in a second step. We present the different types of dislocations and address their positional differences. Modality is considered in greater detail with respect to interrogative and exclamatory sentences. For interrogative sentences, we will introduce a distinction between yes/no interrogatives and wh-interrogatives and highlight the placement of interrogative words in relation to other constituents in the left periphery of the sentence. When considering exclamative sentences, we will differentiate between degree and quantitative exclamatives, as well as the different exclamative words these are associated with.

Keywords: focus, dislocations, interrogatives, exclamatives, syntax, semantics, prosody

1 Introduction

After a brief discussion of the canonical word order in Catalan and a definition of the concept of *information structure*, this chapter explores word order alternations resulting from information structural needs and aspects of modality. More concretely, we will demonstrate the ways in which focus, dislocations, interrogatives and exclamatives have an impact on the canonical word order.

For Catalan, expert consensus seems to be that the order of the verb (V) and its internal arguments (direct object, indirect object or PP object) is fixed: V-DO-IO/PP (see, e.g., Vallduví 1993; 2002, 1230; Wheeler/Yates/Dols 1999; Villalba 2009; López 2009). There is, however, an ongoing debate over the canonical position of the subject; some authors claim that SVO is the normal word order (e.g. Hernanz 2002, 1022; López 2009; Forcadell 2013), while others argue that it is VOS (Vallduví 1992, ch. 5.2; 2002, 1245; Rosselló 2000). Wheeler/Yates/Dols (1999) assume an intermediate

<https://doi.org/10.1515/9783110450408-009>

position, in which the postverbal subject is compulsory only with unaccusative verbs (*Venen trens* ‘Trains are coming’; see 75.2 The Simple Sentence for details on intransitive verbs), while being optional in all other cases (depending on informativeness, phonological weight, etc.).

In this chapter, we assume SVO to be the canonical word order in Catalan. One important argument in support of this view stems from so-called all-new contexts (López 2009, 132; Forcadell 2013). In all-new contexts, the entire utterance is considered to be new, non-given information that appropriately answers the question “What happened?”.

(1) Context: What happened?

- a. La nena va robar el llibre.
 The girl PST.3SG steal the book
 ‘The girl stole the book.’
- b. *Va robar el llibre la nena.
 PST.3SG steal the book the girl

In (1a), the subject is in preverbal position and as such constitutes an appropriate answer to the context question. In (1b), in contrast, the subject is sentence final. This order is not plausible for an all-new reading. As a consequence, VOS cannot be taken as the canonical word order in Catalan (see Forcadell 2013 for a detailed study employing this line of reasoning).

Even though the propositional content of the two utterances in (1) is equivalent, their interpretation is not; while (1a) is appropriate in the context of “What happened?”, (1b) is not. (1b) would be suitable in a context such as “Who stole the book?”, while (1a) would not – at least not with the same intonation as in the “What happened?” context. The structuring of a sentence by syntactic, prosodic or morphological means that correspond to the communicative demands of a specific context or discourse is referred to as the sentence’s *information structure* or *information packaging* (Vallduví/Engdahl 1996, 460; Féry/Ishihara 2016). These means provide cues for the addressee to correctly interpret the meaning intended by the speaker. The addressee should be able to identify which part of the sentence is an actual contribution to its information state at the time of the utterance and which part represents material already subsumed by this information state (Vallduví/Engdahl 1996, 469; Feldhausen 2010, 9). We will take a closer look at these two aspects in sections 2 and 3.

1.1 Sentence modality

Sentence modality, also referred to as sentence force or sentence type, is a formal mark of the illocutionary force of an utterance, namely a conventional grammatical

encoding of the use an utterance has as speech act. For instance, we usually ask for information with interrogative sentences, or request that others do something by means of imperative sentences.

This correspondence is not absolute, however. On the one hand, both (2) and (3) involve interrogative sentences, but only (2) is a genuine question speech act:

- (2) Què vas fer ahir?
 what PST.2SG do yesterday
 ‘What did you do yesterday?’
- (3) Qui es podia imaginar això?
 who REFL can.PST.3SG imagine this
 ‘Who could even imagine this?’ [= ‘Nobody could imagine this.’]

Whereas the interrogative sentence in (2) works as a demand for information, the value of the interrogative sentence in (3) is equivalent to an assertion. Traditionally, utterances like (3) are called rhetorical questions, since they do not actually seek any answer, but rather make a statement in an indirect manner.

Similarly, even though both (4) and (5) contain exclamative sentences, only (4) would count as an expressive speech act (an exclamation), as subordinate sentences by definition lack illocutionary force:

- (4) Que alta que és!
 how tall.F that is
 ‘How tall she is!’
- (5) No t’imagines com és d’alta.
 not REFL-imagine.2SG how is of-tall.F
 ‘You can’t imagine how tall she is.’

In this chapter, we will study the main formal properties of the interrogative and exclamative sentence modalities, which prototypically encode questions and exclamations, respectively. This survey, however, is not an exhaustive exploration of the full range of associated linguistic features, which would involve looking at intonational and pragmatic aspects. Hence, the reader is referred to the following literature for a broader perspective. Escandell-Vidal (2012) and Kissine (2013) are good recent introductions to speech acts, in addition to classical works by Searle (Searle 1969; 1979). As far as prosody and intonation are concerned, the reader can consult Prieto (2002), Prieto et al. (2015), and Font Rotchés (2007). Interrogative sentences are described at length by Brucart/Rigau (2002), while Villalba (2008) is a good source for exploring exclamatives.

The chapter is divided into two main parts. In the first part, we will describe how Catalan packages information and how this packaging affects syntax and prosody, concentrating on focus (section 2) and dislocation structures (section 3). In the second

part, we will describe the interrogative (section 4) and exclamative modalities (section 5), stressing the parallels and differences between these with regards to word order, as well as providing a list of particular interrogative and exclamative words and constructions.

2 Focus

Focus can be understood as the part of the sentence which makes an actual contribution to the information state of the hearer at the time of the utterance (Vallduví/Engdahl 1996). The material providing the actual new information can consist of different constituents within an utterance – depending on the context (6).

- (6) Context: What happened?
- a. [La nena va robar el llibre]_{F(OCUS)} all-new focus
 The girl PST.3SG steal the book
 ‘The girl stole the book.’
- Context: What did the girl do?
- b. La nena [va robar el llibre]_{F.} broad focus
 The girl PST.3SG steal the book
- Context: What did the girl steal?
- c. La nena va robar [el llibre]_{F.} narrow focus
 The girl PST.3SG steal the book

In (6a), the entire sentence constitutes an actual contribution to the hearer’s information state, and hence adheres to the notion of all-new focus. In (6b), the verb and the object are new, while the subject is contextually given (broad focus). In (6c), only the object contributes to the information state (narrow focus). As indicated by the context question, the actual contribution can be considered to be neutral information. The person asking lacks a certain piece of information and the person responding contributes his or her knowledge to the information state of the first person. Focus can also be contrastive (see, e.g., Vallduví 2003 and Repp 2016, for details on contrast). In (7), the first speaker believes that the girl stole the pen, but the second speaker corrects him or her by saying that it was in fact the book that the girl stole. Thus, book is contrasted with pen (CF=contrastive focus):

- (7) Context: Did the girl steal the pen?
- (No.) La nena va robar [el llibre]_{CF}.narrow focus
 No. The girl PST.3SG steal the book

In the last few decades, several studies have investigated realization strategies of (neutral and contrastive) focus in Catalan (e.g. Vallduví 1993; Estebas-Vilaplana 2000; Domínguez 2002; López 2009; Vanrell/Fernández-Soriano 2013; Forcadell 2013; 2016; Feldhausen/Vanrell 2014). In his seminal work on Central Catalan, Vallduví (1992; 1993) proposes an analysis which can easily be described as the standard approach to focus in Catalan. As detailed above, focus can be realized by syntactic, prosodic and/or morphological means, and languages differ in their strategies for realizing focus (see Vallduví/Engdahl 1996; Féry/Ishihara 2016, among many others). Vallduví (1993) claims that Catalan relies solely on syntactic means for the representation of focus, while the position of sentential stress is invariant and is located at the rightmost position of the core clause (Vallduví 1992, 472; 1993, 110; see also Vanrell/Fernández-Soriano 2013). Focus is realized by moving all non-focal constituents out of the core clause and “whatever is left in the core clause (under the lowest IP [= Inflection Phrase]) must be interpreted as focal (with the exception of clitics)” (Vallduví 1993, 119). In the case of all-new sentences, no element is dislocated (see (12); small capitals indicate main stress), while in broad focus, the given material is (right- or left-) dislocated (such as the subject *la nena* in (6b) or the PP object *al calaix* in (9)). For narrow focus, everything but the focused element is dislocated: (a) In (10), where the verb is focal, the two objects are (right- or left-) dislocated; (b) if a constituent other than the verb is narrowly focused, such as the direct object *molts amics* in (11), everything else is right-dislocated, including the verb. Note that the word order in (11a) is traditionally analyzed as focus preposing (e.g. Bonet/Solà 1986, 138–139; López 2009). Vallduví, however, reanalyzes this structure as a right-dislocation (see (11b)), and thus proposes the same syntactic structure for right-dislocation and focus preposing. One argument stems from the fact that the linear order of the post-focal material is free – exactly as is known to be the case for right-dislocations; cf. (11b) vs. (11c); see Vallduví (1992; 1993, 119–131) for further details. Examples (8)–(11) are adapted from Vallduví (1992, 469; 1993, 100–106); square brackets indicate the core clause (i.e. an Inflection Phrase); *t* stands for *trace*.

- (8) [Fiquem el ganivet al CALAIX.]
 put.1PL the knife in.the drawer
 ‘We put the knife in the drawer.’

- (9) a. [Hi₁ fiquem el GANIVET t₁.] al calaix₁.
 CL put.1PL the knife in.the drawer
- b. Al calaix₁, [hi₁ fiquem el GANIVET t₁.]
 in.the drawer CL put.1PL the knife
- (10) a. [L₂’hi₁ FIQUEM] al calaix₁, el ganivet₂
 CL CL put.1PL in.the drawer the knife

- b. El ganivet₂, al calaix₁, [l₂'hi₁ FIQUEM t₂ t₁.]
 the knife in.the drawer CL CL put.1PL
- (11) a. MOLTS AMICS té la Núria.
 many friends has the Núria
 'Núria has many friends.'
- b. [(pro)₁ t₂ MOLTS AMICS], té₂, la Núria₁.
 many friends has the Núria
- c. [(pro)₁ t₂ MOLTS AMICS], la Núria₁, té₂.
 many friends the Núria has

Interestingly, Vallduví's approach does not explain why non-focal, non-verbal elements can always be either right- or left-dislocated (see (13) and (14)), while non-focal, verbal elements must be right-dislocated and cannot appear in the left periphery. His conclusions have been challenged by subsequent studies, such as Estebas-Vilaplana (2000), Domínguez (2002), Vanrell/Fernández-Soriano (2013), Feldhausen/Vanrell (2014) and Forcadell (2016), showing that his ideas might be too restrictive and thus fail to grasp the empirical reality of the different varieties of Catalan. First, there are additional syntactic strategies employed to realize (neutral) focus next to dislocation (such as p-movement or clefting; e.g. Domínguez 2002; Vanrell/Fernández-Soriano 2013; Feldhausen/Vanrell 2014). Second, sentential stress is not invariant, but rather can also be used as a means of realizing focus (e.g. Estebas-Vilaplana 2000; Vanrell/Fernández-Soriano 2013; Forcadell 2016). Finally, the occurrence and the frequency of syntactic and prosodic strategies vary between different varieties of Catalan (Domínguez 2002; Vanrell/Fernández-Soriano 2013; Feldhausen/Vanrell 2014).

Domínguez (2002), for example, argues that Valencian Catalan also employs *prosodically motivated movement* (p-movement; in the sense of Zubizarreta 1998) in order to position the focused element at the right edge of the core clause, (12). In (12a), the PP *al calaix* is moved to the left of the object, with *el ganivet* thus constituting the unambiguous narrow focus of the utterance; marked by [...]_F. In Central Catalan, in contrast, the PP object must be dislocated and cannot be p-moved. Narrow focus on the subject can be easily explained by assuming p-movement in Valencian Catalan: The material following the subject is p-moved to a position before the subject (*el Pep* in (12b)). The possible order of the constituents within a core clause is less restricted in Domínguez's account than in Vallduví's, with non-focal material being allowed to remain within the core clause.

- (12) a. [Fiquem al calaix_i [el GANIVET]_{F t_i}]
 put.1PL in.the drawer the knife
 ‘We put the knife in the drawer.’
- b. [Ficarà el ganivet_i al calaix_j [el PEP]<sub>F t_i t_j]
 put.FUT.3SG the knife in.the drawer the Pep
 ‘Pep will put the knife in the drawer’</sub>

Dialectal variation is also addressed in Vanrell/Fernández-Soriano (2013) and Feldhausen/Vanrell (2014). These works experimentally demonstrate that strategies such as fronting (13), right- and left-dislocation (see above), clefting (14), and in situ realization (15) are used for neutral focus in Balearic, Central, and Valencian Catalan. Additional evidence for in situ realization stems from Forcadell (2016). Based on a corpus study of Catalan TV3 series, she shows that patterns such as that seen in (15) exist in Central Catalan. She argues at length that these patterns cannot be treated as right-dislocations with clitic-drop (Forcadell 2016, ch. 4.1).

- (13) Un MERCEDES s’ha comprat el Jordi.
 a Mercedes REFL-have.3SG bought the Jordi.
 ‘Jordi has bought a Mercedes.’
- (14) Va ser la MARIA la que va portar el cotxe a la seva
 cosina.
 PST.3SG be the Maria the that PST.3SG carry the car to the her
 cousin
 ‘It was Maria who brought the car to her cousin.’
- (15) [FIQUEM el ganivet al calaix.]
 put.1PL the knife in.the drawer
 ‘We put the knife in the drawer.’

These studies show that a clear division between word order (i.e. syntax) and intonation focal typology may be too rigid, since the “word order language” Catalan allows for both mechanisms to different degrees. Vanrell/Fernández-Soriano (2013) further demonstrate that the different strategies and frequencies also occur in interrogatives as well as yes-no and wh-questions.

Work by Vanrell/Fernández-Soriano (2013), Sánchez Candela (2013), and Feldhausen/Vanrell (2014) additionally reveals that syntactic strategies such as clefting and focus fronting are not limited to contrastive contexts as previously assumed (see, e.g., López 2009). If structures such as focus fronting, clefting or in situ realization are also used for neutral focus, the question arises as to how speakers can tell the difference between the different focus types. First, a clear frequency dependent

difference seems to exist (see Vanrell/Fernández-Soriano 2013; Feldhausen/Vanrell 2014). Clefting, for example, is the most frequent strategy for realizing contrastive focus, but is subject to considerable limitations in its use for neutral focus (if used at all, only subject constituents can be clefted; Feldhausen/Vanrell 2014, 123). Second, there seem to be prosodic differences that distinguish between the two types of focus, and hearers can clearly perceive these differences (see Estebas-Vilaplana 2000; Vanrell/Fernández-Soriano 2013; Vanrell et al. 2013; Prieto et al. 2015).

Finally, as the classical approach by Vallduví (1993) appears to have inherent limitations pertaining to dialectal variation as well as variation in the strategies used to realize focus, the challenge becomes defining an approach that accounts for attested variation. To our knowledge, no such formal approach exists thus far. López (2009) and Feldhausen/Vanrell (2014) have proposed updated strategies, but fail to account for the whole picture, which includes neutral and contrastive focus, the frequency of different syntactic and intonational strategies, as well as dialectal variation.

3 Dislocation

Catalan, as has previously been mentioned, exhibits a phenomenon known as dislocation. Dislocations are characterized by the presence of a phrase either in the first position of the clause or at the end of the clause. The dislocated constituent is connected with the clause by means of a resumptive element.

- (16) Normal word order (V-DO-PP):

Va portar les cadires al pis.
 PST.3SG bring the chairs to-the flat
 'She/He brought the chairs to the flat.'

- (17) Clitic Left-Dislocation (CLLD):

Les cadires_i, les_i va portar al pis.
 the chairs CL.ACC PST.3SG bring to-the flat
 'She/He brought the chairs to the flat.'

- (18) Clitic Right-Dislocation (CLRD):

Les_i va portar al pis, les cadires_i.
 CL.ACC PST.3SG bring to-the flat the chairs
 'She/He brought the chairs to the flat.'

- (19) Hanging Topic Left-Dislocation (HTLD):

Les cadires_i, (dius?) la Carme en_i va parlar ahir.
 the chairs (say.2SG) the Carme CL.PART PST.3SG speak yesterday
 'Carme talked about the chairs yesterday.'

While the determiner phrase (DP) *les cadires* ‘the chairs’ is found in its canonical position in (16), it is left-dislocated in (17) and right-dislocated in (18). In both dislocation structures, the dislocated DP is doubled by the clitic pronoun, which agrees in case, gender and number with the dislocated phrase. As the resumptive element is a clitic, this structure is called a clitic left- or right-dislocation.

A second type of dislocation – illustrated in (19) – is known as a hanging topic left-dislocation (see Cinque 1977; Villalba 2009; López 2016), which can be structurally distinguished from the former by the type of resumptive element in the core clause. In (19) it is a weak pronoun, namely *en*, but it can also be an epithet (such as *baluernes* in *Les cadires, la Carme va parlar ahir d’aquestes baluernes* ‘The chairs, Carme spoke of these lumbers yesterday’) or a strong pronoun (such as *elles* in *Les seves germanes, la Carme va anar a la platja sense elles* ‘Her sisters, Carme went to the beach without them’). Further differences between HTLD and CLLD/CLRD exist with respect to their syntax and information structure, but due to space limitations, we will not deal any further with HTLD here (for further information see Villalba 2009, sec. 2.2; López 2016), nor do we address the under-researched area of hanging topics dislocated to the right of the core clause (see López 2016 for information).

Different maximal projections can be clitic left- and right-dislocated in Catalan: PPs (*De la Maria, en vam parlar ahir* ‘We talked about Maria yesterday’), APs (*D’intel·ligent, no ho és pas* ‘S/he is not intelligent’), AdvPs (*Obertament, la Maria no hi ha parlat mai* ‘Maria has never talked overtly’), and CPs (*Que té por, ho sap tothom* ‘Everybody knows that s/he is afraid’); see Villalba (2009, 45–46, 99–100). Even verbal projections can be dislocated (e.g. (11b,c), see Vallduví 2002, sec. 4.6.1 for details). Direct objects, indirect objects and locative arguments are always resumed by a clitic (Vallduví 2002, 1233–1236). Since there are no nominative clitics, dislocated subjects are not resumed by a clitic; it is typically assumed that subject agreement takes over the resumptive function (Vallduví 2002, 1242; López 2016). For dislocated adjuncts, the corresponding resumptive pronouns are optional (Vallduví 2002, 1261–1262). The order of clitics in Catalan is DAT-ACC-PART(itive)-LOC(ative), even though a good deal of syncretism and dialectal variation also occurs (cf. Bonet 2002, 973). Clitic left- and right-dislocations are not unique to Catalan, but also appear in other Romance and non-Romance languages (see, e.g., Villalba 2009; López 2016). However, Catalan is known to make greater use of dislocations than, for example, English or Spanish (Vallduví 1992; Villalba 2011; Feldhausen/Vanrell 2014).

From a prosodic point of view, CLLDs are accented and end with a continuation rise (Prieto 2002, 411; Feldhausen 2010, ch. 5). The prosodic boundary at the right edge is obligatory and can be located at the “intermediate phrase” or the intonational phrase level (Feldhausen 2010, ch. 5). Left-dislocations in subordinated clauses typically phrase with the matrix clause and are not separated by a prosodic boundary at their left edge (Feldhausen 2010, 162). This is illustrated in (20). Here, the subordinated clause “[_{CP2} ...]” is embedded in the matrix clause “[_{CP1} ...]”. As can be seen, the sentence-internal prosodic boundary, indicated by “)(“, does not align with the left

[±a] and contrast [±c] instead – which should lead to clearer, falsifiable predictions (López 2009, 71). A left-dislocation is [+a] and [+c], whereas a right-dislocation is only [+a]. López distinguishes CLLD from CLRD solely by their value for contrastiveness, and not by their being either a link or a tail. Like Villalba, López (2009, 54) criticizes the inherent notion of a tail. While contrastiveness is a basic concept in López's account, Vallduví and Villalba consider contrastiveness to be an epiphenomenon of CLLD (Vallduví 1993, 89; Villalba 2009, 65–66). A detailed investigation of the interpretative properties of CLLD and CLRD is presented in Villalba (2009, 56–67, 108–115) and López (2009, ch. 2). Further details on the information structural interpretation of CLLD and CLRD can be found in Bartra (1985), Villalba (2011), Vallduví/Engdahl (1996), and Leonetti (2011).

As described before, CLLD is typically assumed to be contrastive. However, evidence exists suggesting that CLLD is not necessarily contrastive, such as (22), for example. Imagine a situation in which a couple comes home from shopping, weighed down by many bags. Both enter their apartment with the goals of putting the bags down, going to the toilet, and putting away their purchases. After one of the two returns from the bathroom, she/he utters the sentences in (22). The left-dislocation *la porta* 'the door' is salient, since the couple recently entered their apartment. This is not, however, contrastive. There is no appropriate alternative for *la porta* that makes sense. The windows of the flat or the doors of the refrigerator, for example, are irrelevant. Putting away the purchases is not an alternative either, because this would contrast with the whole proposition and not solely with the left-dislocation.

- (22) Escolta, ara que me'n recordo. La porta, l'has tancada?
listen now that me.CL recall the door CL.ACC.have.2SG closed.F
'Listen, now that I remember. Did you close the door?'

- (23) L'has tancada, la porta?
CL.ACC.have.2SG closed.F the door
'Have you CLOSED the door?'

Nevertheless, instead of using a left-dislocation, partner A could have also used a right dislocation (23). The fact that right-dislocations are very common in Catalan (Villalba 2011) and cannot be associated with a contrastive reading might explain why this position is typically used for non-contrastive material. The mere existence of such a position at the right edge of the sentence does not mean that a left-dislocation is automatically contrastive, however, as is shown in the example above.

4 Interrogative sentences

We typically categorize interrogative sentences according to the type of information requested. Hence, yes/no interrogatives inquire about the truth of a whole proposition (24a), whereas partial interrogatives draw attention to a particular constituent of the sentence, i.e. the *wh*-word (24b) (GIEC 2016, sec. 34.2):

- (24) a. Voleu un cafè?
 want.2PL a coffee
 ‘Do you want a cup of coffee?’
 b. Què et van dir?
 what ACC.CL PST.3PL say
 ‘What did they say to you?’

Yes/no interrogatives such as (24a) can only be distinguished from declarative sentences by means of the typical final rising intonation (cf. Prieto et al. 2015, 21 for details on dialectal variation), even though some grammatical marks are also possible, as discussed in sections 4.1 and 4.2.

It must be emphasized that in the case of partial interrogatives, the interrogative word in the question correlates with the focus constituent of the answer (see section 2):

- (25) a. Què voleu?
 what want.2PL
 ‘What do you want?’
 b. (Volem) un café.
 want.1PL a coffee
 ‘We want a cup of coffee.’

In this section, we will consider two major syntactic aspects of interrogative sentences: the repertory of specialized interrogative words and the word order typically associated with the interrogative modality. Since Catalan is not particularly distinct from other Romance languages in its use of these syntactic aspects, grammatical descriptions have been quite hard to come by (Brucart/Rigau 2002, sec. 8.2.4.1; GIEC 2016, sec. 34.2; Ordóñez 2007). In contrast, there is a growing body of literature available on intonation (see Font Rotchés 2007; Prieto 2002; Prieto et al. 2015) and the syntax-prosody interface (Planas-Morales/Villalba 2013; Prieto/Rigau 2011; Villalba/Planas-Morales 2016). In the following, we will concentrate on the description of the syntactic core of the interrogative system in Catalan.

- (29) No sé si la Carme vindrà.
 not know.1SG whether the Carme come.FUT.3SG
 ‘I don’t know whether Carme will come.’

4.2 Embedded interrogative sentences

In Catalan, we distinguish between two main types of embedded interrogative sentences. On the one hand, the interrogative subordinate clause may reproduce a question speech act (see Villalba 2002, sec. 18.5.2.2). Hence, the yes/no question in (30a) and the partial question in (31a) can be reported as (30b) and (31b), respectively:

- (30)a. El Jordi va preguntar: “Teniu gana?”
 the Jordi PST.3SG ask have.2PL hunger
 ‘Jordi asked: “Are you hungry?”’
 b. El Jordi va preguntar que si teníem gana.
 the Jordi PST.3SG ask that whether have.SBJ.2PL hunger
 ‘Jordi asked whether we were hungry.’
- (31) a. El Jordi va preguntar: “Qui té gana?”
 the Jordi PST.3SG ask who has hunger
 ‘Jordi asked: “Who’s hungry?”’
 b. El Jordi va preguntar que qui tenia gana.
 the Jordi PST.3SG ask that who had hunger
 ‘Jordi asked who was hungry.’

Note that the embedded sentence is introduced by means of the complementizer *que* ‘that’, which appears whenever we report direct speech:

- (32) a. El Jordi va dir: “Calleu!”
 the Jordi PST.3SG say shut.up.SBJ.2PL
 ‘Jordi said: “Shut up!”’
 b. El Jordi va dir que calléssim.
 the Jordi PST.3SG say that shut.up.SBJ.1PL
 ‘Jordi told us to shut up.’

Since this construction is reporting a real question speech act, the embedded sentence must be selected by a verb capable of introducing direct speech acts. Henceforth, we find (i) pure interrogative verbs like *preguntar* ‘to ask’ (30)-(31), (ii) verbs of saying like *dir* ‘to say’ (32), and (iii) verbs of manner of speaking like *cridar* ‘to shout’:

- (33) a. El Jordi va cridar: “Què dimonis feu?”
 the Jordi PST.3SG say what devils do.2PL
 ‘Jordi shouted: “What the hell are you doing?”’
- b. El Jordi va cridar que què dimonis féiem.
 the Jordi PST.3SG say that what devils do.SBJ.1PL
 ‘Jordi shouted (that) what the hell we were doing.’

On the other hand, some verbs not introducing direct speech do select embedded interrogative sentences, which cannot be considered a reproduction of a question speech act (note the impossibility of the complementizer *que* ‘that’):

- (34) a. El Jordi va veure/descobrir (*que) què havíem fet.
 the Jordi PST.3SG see/discover that what have.PST.1PL done
 ‘Jordi saw/discovered what we had done.’
- b. El Jordi va anunciar/confirmar (*que) què volíem.
 the Jordi PST.3SG announced/confirmed that what want.PST.1PL
 ‘Jordi announced/confirmed what we wanted.’

Indirect interrogatives are selected by a wider range of verbs than reproduced questions, with two big subclasses: intensional interrogative verbs (35), which select an indirect interrogative interpreted as a question, and extensional interrogative verbs (36), which select an indirect interrogative interpreted as an assertion. See Villalba (2002, sec. 18.5.2.2) for a detailed typology.

- (35) La Núria {va preguntar / no sabia} què passava.
 the Núria PST.3SG ask /not knew.3SG what happened
 ‘Núria asked/didn’t know what was going on.’
- (36) La Núria {va dir /va recordar} què passava.
 the Núria PST.3SG ask / PST.3SG remember what happened
 ‘Núria said/remembered what was going on.’

Indirect interrogatives with a *wh*-word may freely alternate with free relatives without any change in meaning:

- (37) La Núria {va preguntar / no sabia} el que passava.
 the Núria PST.3SG ask /not knew.3SG the that happened
 ‘Núria asked/didn’t know what was going on.’
- (38) La Núria {va dir /va recordar} el que passava.
 the Núria PST.3SG ask /PST.3SG remember the that happened
 ‘Núria said/remembered what was going on.’

As the reader may have noticed, verbs like *preguntar* ‘to ask’ and *dir* ‘to say’ select for both kinds of interrogative clauses, so we can obtain the following triple:

- (39) a. La Núria {va preguntar /va dir}: “Què passa?”
 the Núria PST.3SG ask /PST.3SG say what happens
 ‘Núria asked/said: “What’s going on?”’
 b. La Núria {va preguntar /va dir} que què passava.
 the Núria PST.3SG ask /PST.3SG say that what happened
 ‘Núria asked/said what was going on.’
 c. La Núria {va preguntar / va dir} què passava.
 the Núria PST.3SG ask /PST.3SG say what happened
 ‘Núria asked/said what was going on.’

(39a) is an instance of a direct question, (39b), a reported version of the direct question, and (39c) an indirect interrogative.

4.3 Movement and inversion

As is common in Indo-European languages, interrogative words in Catalan are placed by default in the left periphery of the sentence, between the main verb and dislocates (Bartra 1985; Villalba 2002; 2009):

- (40) a. La capsas, on la té, (la Joana)?
 the box where CL.ACC have.2SG the Joana
 ‘The box, where does Joana have it?’
 b. A València què hi visites (sempre)?
 to València what CL.LOC visit.2SG always
 ‘In Valencia, what do you always visit?’

Crucially, neither the subject nor any preverbal adverb may appear between the *wh*-word and the verb; only negation and clitics are allowed (GIEC 2016, sec. 34.2.3; Planas-Morales/Villalba 2013).

- (41) a. La capsas, on (*la Joana) la té?
 the box where the Joana CL.ACC have.2SG
 ‘The box, where does Joana have it?’
 b. A València què (*sempre) hi visites?
 to Valencia what always CL.LOC visit.2SG
 ‘In Valencia, what do you always visit?’

Indirect interrogative clauses display the exact same pattern:

- (42) No recordo la capsa, qui la volia.
 not remember.1SG the box who CL.ACC wanted
 ‘I don’t remember who wanted the box.’

This adjacency condition between the *wh*-word and the verb is less strict with *si* ‘whether’ and *per què* ‘why’ (Rigau 1984; Villalba 2002):

- (43) a. No sé la capsa si/per què la volia.
 not know.1SG the box whether/for what CL.ACC wanted
 ‘I don’t remember whether/why (s)he wanted the box.’
 b. No sé si/?per què la capsa la volia.
 not know.1SG whether/for what the box CL.ACC wanted
 ‘I don’t remember whether/why (s)he wanted the box.’

The possibility of having an interrogative word in situ is restricted to echo-questions, as in the following example:

- (44) Que va venir qui?!
 that PST.3SG come who
 ‘(You say) that who came?!’

These interrogatives cannot be neutral, but are rather marked, demanding clarification or expressing surprise (see GIEC 2016, sec. 34.2.4.1).

5 Exclamatory sentences

In this section, we will study the main formal properties of the exclamative sentence type, which prototypically encodes exclamations. The formal diversity of exclamative sentences and their interface position halfway between syntax, semantics and pragmatics has been a major challenge to any attempt to offer a clear-cut definition. Indeed, most efforts have been devoted to arguing that exclamatives are semantically different from interrogatives. In this enterprise, Elliott (1971) deserves credit for designing a series of tests to determine the exclamative character of English sentences in contrast to interrogative sentences. However, since then and despite much effort, we have not achieved a clear consensus on what defines the exclamative sentence type, besides the recognition of certain “ingredients” (see Villalba 2008 for a summary of the main issues). In the following section, we will present the main features that help identify prototypical exclamative sentences in Catalan, namely the set of exclamative words and the particular word order associated with this modality.

5.1 Exclamative words

For Catalan, it is a common practice among grammarians to distinguish between *qualitative* and *quantitative exclamatives* (Brucart/Rigau 2002, sec. 8.2.4.2; GIEC 2016, sec. 34.3; Villalba 2008). Qualitative exclamatives quantify over degrees of a property, as in the following examples, in which the speaker expresses surprise over the cost of the wine (45a) or the special qualities of the music that the hearer is listening to (45b) (see Castroviejo 2008; Villalba 2008).

- (45) a. Que bo (que) és aquest vi!
 what good that is this wine
 ‘How good this wine is!’
 b. Quina musica (que) escoltes!
 which music that listen.2SG
 ‘What music you are listening to!’

The exclamative word is optionally followed by the complementizer *que* ‘that’. As discussed in Villalba (2016), even though including *que* is the most common solution in Catalan, it is a quite recent phenomenon in the language’s history that was first documented in the nineteenth century.

An alternative to the exclamative of degree in (45a) involves the exclamative word *com* ‘how’ modifying an adjective in situ (Villalba 2003):

- (46) Com és de bo aquest vi!
 what is of good this wine
 ‘How good this wine is!’

Synonymous alternatives to (46) exist with the complementizer-like particle *que* ‘that’ or *si* ‘so’ (see Solà 1999; GIEC 2016, sec. 34.3; Villalba 2003):

- (47) Que/Si n’és de bo, aquest vi!
 that/so CL.PART-is of good this wine
 ‘How good this wine is!’

In contrast to the previous qualitative exclamatives, quantitative ones quantify over the cardinality of a set (the number of books in (48a)), the degree of intensity of an action (the time spent dancing in (48b)), or the degree to which one holds a property or is in a certain state (the amount of patience in (48c)):

- (48) a. Quants llibres (que) llegeixes!
 how.many.PL books that read.2SG
 ‘How many books you read!’

- b. Com hi vam ballar, a la festa!¹
 how CL.LOC PST.1PL dance to the party
 'How much we danced at the party!'
- c. Quina paciència (que) es necessita!
 which patience that REFL need.3SG
 'How much patience is needed!'

Before leaving *wh*-exclamatives, one should note the fact – originally noted by Elliott (1971) for English – that certain *wh*-words are possible in embedded exclamatives, but impossible in root contexts. Therefore, while the indirect exclamatives in (49) are perfect, their corresponding matrix exclamatives are impossible (see Castroviejo 2008 for discussion):

- (49)a. *(Mira) qui va comprar un llibre!
 look.IMP.2SG who PST.3SG buy a book
 'Look who bought a book!'
- b. *(No t'imagines) on va anar.
 not REFL-imagine.2SG where PST.3SG go
 'You can't imagine where (s)he went.'

In addition to *wh*-exclamatives, nominal exclamatives are also very common (GIEC 2016, sec. 34.3.1.4; Villalba 2008):

- (50)a. Les idees que té, aquesta noia!
 the ideas that has this girl
 'What ideas this girl has!'/ 'How many ideas this girl has!'
- b. No t'imagines la Núria les coses que fa.
 not REFL-imagine the Núria the things that does
 'You can't imagine the things this girl does!' [=how many/such things]

Depending on the context, definite exclamatives can have both qualitative and quantitative readings, as the translations of (50) show. Note also that the definite phrase follows a dislocate subject in (50b) (on dislocation, see section 3), and unlike *wh*-exclamatives, it requires the presence of the complementizer *que* 'that'.

¹ This sentence admits a second reading as a qualitative exclamative quantifying over the surprising way that we danced: 'How we danced at the party!'

5.2 Movement and inversion

As is the case with interrogatives (see section 4.3), Catalan shows obligatory subject-verb inversion with exclamatives (Villalba 2017):

- (51) a. Que fàcil que és l'examen!
 how easy that is the.exam
 'How easy the exam is!'
 a'. *Que fàcil que l'examen és!
 how easy that the-exam is
 b. Les/Quines coses que fa la Núria.
 the/which things that does the Núria
 'The/What things Núria does!'
 b'. *Les/Quines coses que la Núria fa.
 the/which things that the Núria does

Yet, the condition of adjacency between the verb and the exclamative word, which forbids preverbal subjects, is less strict with preverbal adverbs:

- (52) a. ?Que fàcils que sempre són els exàmens!
 how easy that always be.3PL the exams
 'How easy the exams always are!'
 b. Les/Quines coses que sovint fa la Núria.
 the/which things that often does the Núria
 'The/What things Núria often does!'

In any event, the inversion pattern (also found in Spanish) is not widespread among Romance languages; French, Italian or Portuguese tend to favor or force the subject-verb word order (Villalba 2017).

6 Conclusions

In this chapter, we have reviewed the main features of information structure and sentence modality in Catalan. Concerning information structure, we have shown that Catalan typically places informative/neutral focus at the rightmost position in the sentence, where it receives the main stress. Nevertheless, experimental studies have shown that additional strategies such as *in situ* realization can also be employed, and that there is dialectal variation in the strategies used for focus marking. Contrastive focus can be obtained *in situ* as well, or in a left peripheral position by means of focus preposing. Catalan thus makes use of word order and intonation to different degrees, yielding a rather complex typology of focus phenomena. Ground material is anapho-

ric and typically appears dislocated to either the left or the right. Clitic left-dislocation is typically contrastive, though not inherently so, whereas clitic right-dislocations usually mark background material that the speaker wishes to make salient again.

Regarding sentence modality, we have shown that both interrogatives and exclamatives involve obligatory subject-verb inversion. Exclamatives constitute a major departure from most Romance languages, which prefer the non-inverted word order. Finally, we have shown that the range of interrogative and exclamative words is very wide, exhibiting partial overlaps, particularly in subordinate contexts.

7 Bibliography

- Astruc, Lluïsa (2005), *The Intonation of Extra-Sentential Elements in Catalan and English*, PhD thesis, Cambridge, University of Cambridge.
- Bartra, Anna (1985), *Qüestions de la sintaxi d'ordre en català*, PhD thesis, Barcelona, Universitat Autònoma de Barcelona.
- Bonet, Eulàlia (2002), *Cliticització*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 1, Barcelona, Empúries, 933–989.
- Bonet, Sebastià/Solà, Joan (1986), *Sintaxi generativa catalana*, Barcelona, Enciclopèdia Catalana.
- Brucart, José María/Rigau, Gemma (2002), *La quantificació*, in: Joan Solà et al. (edd.), *Gramàtica del Català Contemporani*, vol. 2, Barcelona, Empúries, 1517–1589.
- Castroviejo, Elena (2008), *Deconstructing Exclamations*, *Catalan Journal of Linguistics* 7, 41–90.
- Cinque, Guglielmo (1977), *The Movement Nature of Left Dislocation*, *Linguistic Inquiry* 8, 397–411.
- Domínguez, Laura (2002), *Analyzing Unambiguous Narrow Focus in Catalan*, MIT Working Papers in Linguistics 43, 17–34.
- Elliott, Dale (1971), *The Grammar of Emotive and Exclamatory Sentences in English*, *Ohio State Working Papers in Linguistics* 8, viii–110.
- Escandell-Vidal, Victoria (2012), *Speech Acts*, in: J. Ignacio Hualde/Antxon Olarrea/Erin O'Rourke (edd.), *The Handbook of Hispanic Linguistics*, Oxford, Blackwell, 629–651. DOI:10.1002/9781118228098.ch29.
- Estebas-Vilaplana, Eva (2000), *The Use and Realisation of Accentual Focus in Central Catalan with a Comparison to English*, PhD thesis, London, University College London.
- Feldhausen, Ingo (2010), *Sentential Form and Prosodic Structure of Catalan*, Amsterdam/Philadelphia, Benjamins.
- Feldhausen, Ingo (2016), *Inter-speaker Variation, Optimality Theory, and the Prosody of Clitic Left-Dislocations in Spanish*, *Probus* 28/2, 293–334.
- Feldhausen, Ingo/Lausecker, Alina (2018), *Diatopic Variation in Prosody: Left- and Right-Dislocations in Spanish*, in: Malte Belz et al. (edd.), *Proceedings of the Conference on Phonetics and Phonology in the German-Speaking Countries (P&P 13)*, Berlin, Humboldt-Universität zu Berlin/Leibniz-Center for General Linguistics (ZAS), 57–60.
- Feldhausen, Ingo/Vanrell, Maria del Mar (2014), *Prosody, Focus and Word Order in Catalan and Spanish: An Optimality Theoretic Approach*, in: Susanne Fuchs et al. (edd.), *Proceedings of the 10th International Seminar on Speech Production (ISSP)*, Cologne, Universität zu Köln, 122–125.
- Féry, Caroline/Ishihara, Shinichiro (edd.) (2016), *The Oxford Handbook of Information Structure*, Oxford, Oxford University Press.
- Font Rotchés, Dolors (2007), *L'entonació del català*, Barcelona, Publicacions de l'Abadia de Montserrat.

- Forcadell, Montserrat (2013), *Subject Informational Status and Word Order, Catalan as an SVO Language*, Journal of Pragmatics 53, 39–63.
- Forcadell, Montserrat (2016), *New Prosodic Patterns in Catalan, Information Status and (De)accentability*, Journal of Pragmatics 97, 1–20.
- GIEC = Institut d'Estudis Catalans (2016), *Gramàtica de la llengua catalana*, Barcelona, Institut d'Estudis Catalans.
- Hernanz, M. Lluïsa (2002), *L'oració*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 2, Barcelona, Empúries, 993–1073.
- Kissine, Mikhail (2013), *From Utterances to Speech Acts*, Cambridge, Cambridge University Press. DOI:10.1017/CBO9780511842191.
- Leonetti, Manuel (2011), *La expresión de la estructura informativa en la sintaxis, un parámetro de variación en las lenguas románicas*, Romanistisches Jahrbuch 61, 338–355.
- López, Luis (2009), *A Derivational Syntax for Information Structure*, Oxford, Oxford University Press.
- López, Luis (2016), *Dislocations and Information Structure*, in: Caroline Féry/Shinichiro Ishihara (edd.), *The Oxford Handbook of Information Structure*. Oxford, Oxford University Press, 402–422.
- Ordóñez, Francisco (2007), *Observacions sobre la posició dels subjectes postverbals en català i castellà*, Caplletra 42, 251–273.
- Planas-Morales, Sílvia/Villalba, Xavier (2013), *The Right Periphery of Interrogatives in Catalan and Spanish, Syntax-Prosody Interactions*, Catalan Journal of Linguistics 12, 193–217.
- Prieto, Pilar (2002), *Entonació*, in: Joan Solà et al. (edd.), *Gramàtica del Català Contemporani*, vol. 1, Barcelona, Empúries, 393–462.
- Prieto, Pilar/Rigau, Gemma (2007), *The Syntax-Prosody Interface: Catalan Interrogative Sentences Headed by “que”*, Journal of Portuguese Linguistics 6/2, 29–59. DOI:10.1093/jae/ejp002.
- Prieto, Pilar/Rigau, Gemma (2011), *Prosody and Pragmatics*, in: Lluís Payrató/Josep M. Cots (edd.) *The Pragmatics of Catalan*, Berlin/Boston, De Gruyter, 17–48. DOI:10.1515/9783110238693.17.
- Prieto, Pilar, et al. (2015), *Intonational Phonology of Catalan and its Dialectal Varieties*, in: Sónia Frota/Pilar Prieto (edd.), *Intonation in Romance*, Oxford, Oxford University Press, 9–62.
- Repp, Sophie (2016), *Contrast, Dissecting an Elusive Information-Structural Notion and its Role in Grammar*, in: Caroline Féry/Shinichiro Ishihara (edd.), *The Oxford Handbook of Information Structure*, Oxford, Oxford University Press, 270–290.
- Rigau, Gemma (1984), *De com “si” no és conjunció i d'altres elements interrogatius*. Estudis Gramaticals 1, 249–278.
- Rosselló, Joana (2000), *A Minimalist Approach to the Null Subject Parameter*, Catalan Working Papers in Linguistics 8, 97–128.
- Sánchez Candela, Noèlia (2013), *És quan dormo que hi veig clar: aproximació a les construccions de clivellament en català*, Llengua i Literatura: Revista Anual de la Societat Catalana de Llengua i Literatura 23, 157–192.
- Searle, John R. (1969), *Speech Acts: an Essay in the Philosophy of Language*, Cambridge, Cambridge University Press.
- Searle, John R. (1979), *Expression and Meaning: Studies in the Theory of Speech Acts*, Cambridge, Cambridge University Press.
- Solà, Joan (1999), *¿Si que és car!*, in: Joan Solà (ed.), *Parlem-ne. Converses lingüístiques*, Barcelona, Proa, 232–234.
- Solà, Joan (2002), *Les subordinades de relatiu*, in: Joan Solà et al. (edd.), *Gramàtica del Català Contemporani*, vol. 1, Barcelona, Empúries, 2455–2565.
- Vallduví, Enric (1992), *Focus Constructions in Catalan*, in: Christiane Laeufer/Terrell A. Morgan (edd.), *Theoretical Analyses in Romance Linguistics*, Amsterdam/Philadelphia, Benjamins, 457–479.
- Vallduví, Enric (1993), *The Informational Component*, PhD thesis, Philadelphia, University of Pennsylvania.

- Vallduví, Enric (2002), *L'oració com a unitat informativa*, in: Joan Solà et al. (edd.), *Gramàtica del Català Contemporani*, vol. 2, Barcelona, Empúries, 1221–1279.
- Vallduví, Enric (2003), *A Theory of Informatics*, in: Javier Gutiérrez-Rexach (ed.), *Semantics: Critical Concepts in Linguistics*, vol. 1, London/New York, Routledge, 359–384.
- Vallduví, Enric/Engdahl, Elisabeth (1996), *The Linguistic Realisation of Information Packaging*, *Linguistics* 34/3, 459–519.
- Vanrell, Maria del Mar/Fernández-Soriano, Olga (2013), *Variation at the Interfaces in Ibero-Romance, Catalan and Spanish Prosody and Word Order*, *Catalan Journal of Linguistics* 12, 253–282.
- Vanrell, Maria del Mar, et al. (2013), *Prosodic Manifestations of the Effort Code in Catalan, Italian and Spanish Contrastive Focus*, *Journal of the International Phonetic Association* 43/2, 195–220.
- Villalba, Xavier (2002), *La subordinació*, in: Joan Solà et al. (edd.), *Gramàtica del Català Contemporani*, vol. 3, Barcelona, Empúries, 2247–2319.
- Villalba, Xavier (2003), *An Exceptional Exclamative Sentence Type in Romance*, *Lingua* 113/8, 713–745. DOI:10.1016/S0024-3841(02)00117-1.
- Villalba, Xavier (2008), *Exclamatives: A Thematic Guide with Many Questions and Few Answers*, *Catalan Journal of Linguistics* 7, 9–40.
- Villalba, Xavier (2009), *The Syntax and Semantics of Dislocations in Catalan: A Study on Asymmetric Syntax at the Peripheries of Sentence*, Cologne, Lambert Academic Publishing.
- Villalba, Xavier (2011), *A Quantitative Comparative Study of Right-Dislocation in Catalan and Spanish*, *Journal of Pragmatics* 43/7, 1946–1961.
- Villalba, Xavier (2016), *L'evolució de les oracions exclamatives-qu de grau en català*, *Caplletra* 60, 211–226. DOI:10.7203/caplletra.60.8454.
- Villalba, Xavier (2017), *Exclamatives, Imperatives, Optatives*, in: Andreas Dufter/Elisabeth Stark (edd.), *Manual of Romance Morphosyntax and Syntax*, Berlin/Boston, De Gruyter, 603–646.
- Villalba, Xavier/Planas-Morales, Sílvia (2016), *Interacció sintàctica i entonativa a les interrogatives catalanes amb dislocació a la dreta*, in: Ana M. Fernández (ed.), *53 reflexiones sobre aspectos de la fonética y otros temas de lingüística*, Barcelona, Laboratori de Fonètica de la Universitat de Barcelona, 311–319.
- Wheeler, Max W./Yates, Alan/Dols, Nicolau (1999), *Catalan: A Comprehensive Grammar*, London, Routledge.
- Zubizarreta, Maria Luisa (1998), *Prosody, Focus, and Word Order*, Cambridge (MA), MIT Press.

Jaume Mateu

5.5 Lexicalized Syntax: Phraseology

Abstract: In this chapter I deal with some basic syntactic and semantic properties of lexicalized phrases. In particular, I concentrate on the distinction between idioms and collocations and apply it to some basic types of lexicalized phrases in Catalan, with a main focus on verbal phrases. Some relevant remarks are also made on lexicalized prepositional, adverbial, and adjectival phrases, among others. Following Nunberg/Sag/Wasow (1994), three different semantic properties are distinguished when dealing with the interpretation of lexicalized phrases: conventionality, compositionality, and transparency. Strictly speaking, only those lexicalized phrases whose meaning is non-compositional can be classified as idioms. In contrast, those expressions that are idiomatic but do have a compositional meaning are better referred to as idiomatic collocations. The semantic property of (non-)compositionality is shown to be correlated with the syntactic property of (in-)flexibility. Furthermore, the conceptual property of transparency/opacity is claimed to be related to the presence/absence of metaphorical motivation (Lakoff 1993; Gibbs 1995; Espinal/Mateu 2010).

Keywords: idioms, collocations, lexicalized phrases, syntactic flexibility, conceptual metaphors

1 Introduction

It is often said that one has an excellent command of a language when one is able to use it in an idiomatic way.* Speaking a language idiomatically typically involves making use of idioms and collocations. To put it in Mel'čuk's (1995) words, the native speaker can be characterized as the person who uses phrasemes, i.e., lexicalized phrases. The main goals of this brief chapter are to offer an overview of some basic types of lexicalized phrases that can be found in Catalan and to provide some understanding on their formation (see Lorente ²2002 for a more comprehensive classification; see also Espinal/Mateu 2005; Ginebra 2000; 2003; 2017; Mestres 2007; and Salvador 1995 for relevant discussion on Catalan phraseology).

The fact that lexicalized phrases are indeed pervasive in any language has led some authors to conclude that free syntax does not typically exist in practice. How-

* I would like to thank the editors for their useful comments and suggestions. Of course, the usual disclaimers apply. This work has been supported by the Spanish grant FFI2017-87140-C4-1-P and the Catalan grant 2017SGR634.

ever, as pointed out by Mendívil (2009), such a conclusion does not seem to be correct. Rather, I will assume along with this author that there appears to be a continuum like the one exemplified in (1), which ranges from typical lexical items (simple words) to free syntactic combinations (free, i.e., non-lexicalized syntax). Assuming that lexicalized expressions are those that are stored in our mental lexicon (see Jackendoff 1997), all but (1g) can be said to belong to this set.

- (1) a. Simple words (*trencar* ‘to break’).
- b. Complex words (derived words like *trencament* ‘breaking’ or compounds like *trencaclosques* ‘puzzle’).
- c. Complex predicates or light verb constructions (*fer trencament* ‘to do breaking’).
- d. Idioms (*trencar-se el cap*, lit. ‘to break one’s head’, i.e., ‘to make a great effort in order to understand something’).
- e. Idiomatic collocations (*trencar el gel* ‘to break the ice’).
- f. Non-idiomatic collocations (*trencar la promesa* ‘to break the promise’).
- g. Free syntax (*trencar el vidre* ‘to break the glass’).

Following a traditional conception of language, it is often assumed that the domain of morphology deals with the formation of simple and complex words (cf. 1a and 1b), whereas phraseology deals with both idioms and collocations (e.g., 1d, e, f). Syntax is said to involve a free combination of words (e.g., 1g). Light verb constructions or multi-word expressions like phrasal verbs (e.g., *tirar endavant* ‘move forward’; cf. Ginebra 2008; Mateu/Rigau 2010) have been claimed to be somewhere in between.

However, the typical modular picture in which morphology and syntax are differentiated and completely separate components of grammar, each one having its own set of primitives and principles of combination, does not seem to accord with some recent syntactic theories of word-formation. For example, following Hale/Keyser’s (1993; 2002) theory of so-called “lexical syntax”, Espinal/Mateu (2011) claim that the relevant grammatical constraints on the formation of light verb constructions like *fer camí* ‘to do (a) walk’ or *posar en pràctica* ‘to put into practice’ are not so different from the ones involved in denominal verbs such as *caminar* ‘to walk’ or *practicar* ‘to practice’, respectively. In particular, it has been argued that the latter are formed via the syntactic operation of incorporating a nominal head (*camí* ‘walk’ and *pràctica* ‘practice’) into a phonologically null light verb. See Kiparsky (1997) for a critique of Hale/Keyser’s (1993) syntactic theory of denominal verb formation and Espinal/Mateu (2011) for a rebuttal of part of this critique.

Be this as it may, in this brief chapter I will not deal with complex words like (1b) (7.2 Word-Formation) nor, for reasons of space, with light verb constructions like (1c) (see Espinal 2002 and Ginebra 2008; 2017), but rather will focus on an important and interesting distinction put forward in the literature, i.e., that which distinguishes between prototypical idioms (the non-compositional ones) and idiomatic collocations

or idiomatically combining expressions, whose meaning has been claimed to be compositional. Following Mendivil (2009), I prefer to avoid the contradictory label of compositional idioms (cf. Nunberg/Sag/Wasow 1994). Accordingly, if some idioms turn out to be compositional, I will be using his label of “idiomatic collocations”. Indeed, a positive consequence of using this terminology is that it allows us to maintain the standard claim that the meaning of idioms is non-compositional, whereas the meaning of collocations is compositional. Moreover, following Nunberg/Sag/Wasow (1994), I will take pains to show that the semantic property of compositionality must not be confused with that of conventionality. As we will see immediately below, all idiomatic expressions are conventional by definition but some of them are compositional (e.g., 1e) and some of them are not (e.g., 1d). The semantic property of compositionality will be shown to be related to syntactic flexibility. Idioms, which are semantically non-compositional, are predicted to be syntactically inflexible, whereas idiomatic collocations (also known as idiomatically combining expressions or compositional idioms in Nunberg/Sag/Wasow 1994) will be shown to have a more flexible syntax.

2 Idioms and collocations

As noted above, the domain of phraseology includes the study of idioms and collocations. All of these can be characterized as lexicalized phrases. The basic distinction between the former and the latter has to do with semantic compositionality: the former are not compositional, whereas the latter are.

In this section I classify and analyze idioms and collocations according to the category of the head. In Section 2.1, I analyze lexicalized verbal phrases. In Section 2.2, I briefly deal with lexicalized nominal phrases and, finally, in Section 2.3, other types of lexicalized phrases are also summarized.

Before moving forward, one methodological remark is in order: in the linguistic tradition, there has been an emphasis in classifying and establishing different types of lexicalized phrases, in debating their appropriate terminology, in measuring their frequency of use, and in trying to provide full coverage of their types, rather than in understanding the nature of their formation. However, as is stressed by Bosque (2001), providing good classifications or appropriate taxonomies of linguistic constructions (in our case, idioms and collocations) should not be the goal, but rather the starting point of linguistic research. Similarly, this author also claims that, when analyzing the linguistic properties of collocations, frequency of use is not so relevant as the study of the lexical-semantic restrictions involved.

Accordingly, due to limitations of space, my goal in this brief chapter will not be to provide detailed lists or taxonomies of different types of lexicalized phrases in Catalan, but rather to provide some relevant remarks on their basic semantic and syntactic properties.

2.1 Lexicalized verbal phrases

Consider idioms such as (1d) *trencar-se el cap* (lit. ‘to break one’s head’, i.e., ‘to make a great effort in order to understand something’), idiomatic collocations like (1e) *trencar el gel* (‘to break the ice’), and non-idiomatic collocations like (1f) *trencar la promesa* (‘to break the promise’). It is instructive and useful to analyze these three types of lexicalized verbal phrases by taking into account three different dimensions of meaning: conventionality, compositionality, and transparency. Assuming Nunberg/Sag/Wasow’s (1994, 495) definition of conventionality (expressions can be defined as conventional if “their meaning or use can’t be predicted, or at least entirely predicted, on the basis of a knowledge of the independent conventions that determine the use of their constituents when they appear in isolation from one another”), it makes sense to classify idioms like (1d) as strongly conventional. Indeed, in this case both the verb and the selected direct object are used in a clearly idiomatic way, i.e., their meaning or use cannot be predicted in the relevant sense defined above.

Idiomatic collocations like (1e) are also clearly conventional expressions but it is important to realize that in this case it is the direct object *el gel* ‘the ice’ that carries the major idiomatic burden. The figurative meaning of the verb is not so different from the literal one: e.g., the meaning of *trencar el gel* can be claimed to involve *breaking* a social barrier to get something started. Accordingly, idioms like (1d) involve a bidirectional conventionality, whereas idiomatic collocations like (1e) involve a unidirectional one: i.e., in (1e) it is not the case that *trencar* is interpreted as such and such in the idiomatic context of “___*el gel*” (as noted, *trencar* can in fact be interpreted as ‘break’ in (1e)), but rather that *el gel* is interpreted as such and such in the idiomatic context of “*trencar*___”. This is why in (1e) it is the direct object rather than the verb that is said to carry the major idiomatic burden. In contrast, in strongly conventional idioms like (1d) the conventionality relation is said to be bidirectional since *both* the direct object and the verb are affected by the idiomatic burden in parallel proportion. As will be seen immediately below, there is an interesting relation between the conventional directionality type and the semantic property of (non-)compositionality. However, in spite of their relation (e.g., strongly conventional idioms tend to be non-compositional and, conversely, less conventional idiomatic expressions tend to be compositional), I assume along with Nunberg/Sag/Wasow (1994) that these two semantic properties, conventionality and compositionality, should not be confused. Finally, the meaning of non-idiomatic collocations like (1f) can of course be claimed to be much less conventional than the previous idiomatic expressions since its meaning can be predicted “on the basis of a knowledge of the independent conventions that determine the use of their constituents when they appear in isolation from one another” (Nunberg/Sag/Wasow 1994, 495). Thus, conventionality, if understood in this narrow sense, refers to the gradual discrepancy between the figurative or idiomatic reading and the predicted literal meaning of the expression.

A caveat on the notion of conventionality as is generally applied to collocations is in order here: collocations have been described as combinations of words that are preferred over other combinations that otherwise appear to be semantically equivalent. For example, as pointed out by Ginebra/Navarro (2015), although a verbal phrase like *preparar la taula* ‘to prepare the table’ could be understood, the lexicalized phrase expected and used by a native speaker of Catalan would be the collocation *parar taula* ‘to set (the) table’. Typically, as pointed out by Croft/Cruse (2004, 249–250), collocations have been said to be expressions that can be interpreted more or less correctly out of context, but cannot be produced correctly if the conventional expression is not already known to the speech community (cf. also Hausmann 1997, 282).

Next we will see that the semantic property of compositionality allows us to make a different cut when dealing with our lexicalized verbal phrases in (1d,e,f): as noted, the meaning of idioms is non-compositional, whereas the meaning of (non-idiomatic and idiomatic) collocations is compositional. Compositionality refers to the degree to which the phrasal meaning, once known, can be analyzed in terms of how it is distributed among the parts of the expression (cf. Nunberg/Sag/Wasow 1994 and Titone/Connine 1999). For example, idioms like (1d) *trencar-se el cap* (lit. ‘to break one’s head’, i.e., ‘to make a great effort in order to understand something’) are semantically non-compositional since the meaning is *not* distributed among the parts of the idiom: i.e., the expression as a whole is mapped onto the meaning of the idiom. In contrast, both non-idiomatic and idiomatic collocations are semantically compositional since the meaning is distributed among their parts. For example, in idiomatic collocations like (1e) *trencar el gel* (‘to break the ice’), the individual parts of the literal expression can be mapped onto individual parts of the figurative/idiomatic meaning (e.g., *trencar*: ‘break’ // *el gel*: ‘a social tension/barrier...’). Importantly, following Nunberg/Sag/Wasow (1994, 499), I agree with only the weaker claim that speakers are capable of recognizing the compositionality of a phrase like (1e) *trencar el gel* “after the fact, having first divined its meaning on the basis of contextual cues” (emphasis mine: JM). From the previous picture, it can be inferred that idiomatic collocations like (1e) are situated somewhere between non-idiomatic collocations like (1f) and idioms like (1d); in short, (1e) shares semantic compositionality with (1f) and idiomaticity with (1d).

As pointed out above, semantic (non-)compositionality correlates with syntactic (in-)flexibility. Indeed, it is well-known that a typical property of idioms is their syntactic inflexibility. In contrast, collocations are syntactically more flexible. What is expected from the scale of lexicalized phrases exemplified above is that idioms like (1d) *trencar-se el cap* are the least flexible ones, whereas non-idiomatic collocations like (1f) *trencar la promesa* are the most flexible ones. Idiomatic collocations like (1e) *trencar el gel* are more flexible than idioms like (1d) but less flexible than non-idiomatic collocations like (1f). For example, clitic left dislocation constructions are expected to be impossible with idioms, but are possible with collocations: cf. (2a) and (2b,c).

- (2) a. #El cap_i se'l_i va trencar el noi (estudiant).
 the head refl.acc.cl. past.3sg break the boy studying
 b. El gel_i el_i va trencar el noi.
 the ice acc.cl past.3sg break the boy
 c. La promesa_i la_i va trencar el noi.
 the promise acc.cl past.3sg break the boy

The explanation of the contrasts exemplified in (2) is due to the fact that idioms are not semantically compositional, whereas collocations are. The example in (2b) is well-formed since, as noted above, the complex meaning of idiomatic collocations like *trencar el gel* can be distributed among their parts; for example, *el gel* ‘the ice’ can be figuratively related to a particular social tension (see above). Hence this direct object is crucially referential. Notice that a similar distribution is impossible in (2a): *el cap* ‘the head’ is not related to any differentiated or individuated part of the idiomatic meaning (cf. ‘to make a great effort in order to understand something’). The object nominal phrase in an idiom like (2a) is thus non-referential, whereas a discourse referent of the object nominal phrase in an idiomatic collocation like (2b) can be identified. This fact could also be claimed to explain the well-formedness of a pronominal passive construction like *El gel es va trencar amb unes quantes begudes* ‘The ice was broken with a few drinks’. The English example is taken from Ruwet (1991, 173; ex. (5)). In contrast, idioms cannot be passivized: e.g., cf. *L'enemic ha fotut el camp* // **El camp ha estat fotut per l'enemic* ‘The enemy has shoved the camp, i.e., has taken off’ (cf. Ruwet (1991, 173; ex. (2))). It seems reasonable to claim that the impossibility of passivization is related to the non-referential status of the nominal phrase *el camp*.

Furthermore, the following examples in (3) can provide an additional test to distinguish syntactically inflexible idioms from syntactically flexible collocations (cf. Bosque 2001; Mendívil 1999; 2009).

- (3) a. #El noi es va trencar l' enorme cap que tenia (estudiant).
 the boy refl.cl past.3sg break the big head that had studying
 b. ?El noi va trencar el gel que hi havia a l'ambient.
 the boy past.3sg break the ice that there was at the.air
 c. El noi va trencar la promesa que ens va fer.
 the boy past.3sg break the promise that us.dat past.3sg make

Compare the slightly marginal example in (3b) with a fully acceptable non-idiomatic one such as *El noi va trencar la barrera que ens impedia avançar* ‘The boy broke the barrier that prevented us from moving forward’, where the idiomatic object *el gel* ‘the ice’ has been replaced by a non-figurative one like *la barrera* ‘the barrier’. As predicted by Nunberg/Sag/Wasow (1994), to the extent that the idiomatic meaning can be distributed among the parts (e.g., the ice is intended to mean some kind of social

tension), the example in (3b) is predicted to become more acceptable. That is, (3b) can be found to be acceptable insofar as the direct object contributes an *individual* meaning to the idiomatic expression. Accordingly, the modification through a relative clause in (3b) applies to the entity expressed by the referential direct object. In contrast, in the idiom contained in (3a) the modification of the nominal object *cap* ‘head’ is not possible since it is non-referential. Notice that it is not modification in general what is prohibited in (3a). For example, *El noi es va trencar molt el cap estudiant* is perfectly acceptable, since the adverbial quantifier *molt* ‘very much’ does not involve any sort of modification over individuals but conveys an intensive quantification over the whole verbal predicate (see Espinal 2002 for further discussion of this type of modification in the context of idiomatic expressions). Finally, in clear contrast to idioms like (3a), non-idiomatic collocations like (3c) are syntactically flexible without any special restrictions to be noted. Given this, the relevant conclusion seems to be that individual elements of (idiomatic) collocations have (some degree of) semantic autonomy, whereas the individual syntactic elements of idioms exhibit no semantic independence from one another whatsoever.

In relation to the important claim that idioms (also known as “idiomatic phrases”) are non-compositional but idiomatic collocations (also known as “idiomatically combining expressions”) are compositional, it has been noted that the former do not necessarily preserve the aspectual class associated with the literal meaning, whereas the latter do preserve it (see Glasbey 2007; Espinal/Mateu 2010). For example, as shown in (4), *trencar-se el cap* has an atelic interpretation of the idiomatic reading in (4a), whereas *trencar el gel* preserves the expected telicity, i.e., the one associated with the literal reading of *trencar* ‘to break’ plus a definite direct object like *el vidre* ‘the glass’: cf. the aspectual parallelism in (4b) and (4c). See Glasbey (2007) for the proposal that the durativity of the eventuality in (4a) has to do with the absence of a gradual patient, a thematic role that can only be assigned if the object is referential, which is not the case in (4a). However, such a proposal is problematic if we consider other cases such as *fotre el camp* lit. ‘to shove the camp, i.e., leave’, where there is no gradual patient either but the eventuality is still telic (cf. *fotre el camp {a les sis en punt/#durant hores}*). In view of this problem, Espinal/Mateu (2010) claim that the telic-to-atelic event type-shifting involved in (4a) could have a conceptual explanation related to the unbounded nature of the notion of intensity (see below for more discussion).

- (4) a. El noi es trencà el cap {#a les sis en punt/durant hores} per
 saber qui era l’assassí.
 the boy refl.cl broke the head at six o’clock/for hours to
 know who was the.murderer
 b. El noi trencà el gel {a les sis en punt/en dos minuts/#durant hores}
 the boy broke the ice at the six o’clock/in two minutes/for hours

- c. El noi trencà el vidre {a les sis en punt/en dos minuts /#durant hores}
 the boy broke the glass at the six o'clock/in two minutes/for hours

The third semantic property I want to deal with is transparency/opacity, which should not be confused with compositionality/non-compositionality, respectively. As pointed out by Nunberg/Sag/Wasow (1994, 496), an idiomatic expression can be said to be transparent when the speaker can wholly recover the rationale for the figuration it involves. Although it is expected that some correlation can be found between those two semantic properties (e.g., clearly transparent idioms tend to be compositional and, conversely, opaque idiomatic expressions tend to be non-compositional), we will soon see that it is not hard to find transparent idioms that are non-compositional.

According to Lakoff (1993) and Gibbs (1995), among others, the existence of so-called “idiom families” is often due to the presence of conceptual metaphors that motivate their related meanings (i.e., that make their meanings (more) transparent). In particular, Espinal/Mateu (2010) argue that the meaning of idioms like *treure el fetge per la boca* ‘to expel the liver through the mouth’ (cf. Engl. *to work one’s guts out*); *petar-se el cul* ‘to explode the butt’ (cf. Engl. *to laugh one’s butt off*); *sortir-li els ulls de les òrbites* ‘(the eyes) to leave the orbits’ (cf. Engl. *to cry one’s eyes out*), etc. can be shown to be transparent by positing the following underlying conceptual metaphor: AN EXTREME INTENSITY IS AN EXCESSIVE DETACHMENT OF A BODY PART. Their claim is that the meaning of these idioms can be shown to be transparent to the extent that it is motivated by the relevant conceptual metaphor. This notwithstanding, it is important to realize that their meaning is *not* compositional: e.g. the direct object *el cul* ‘the butt’ (cf. *Ens vam petar el cul*; *We laughed our butts off*) is not referential and the meaning of the idiom ‘(to laugh a lot)’ is not distributed onto the parts of the expression. Accordingly, syntactic inflexibility is predicted, as shown by the ill-formed examples in (5): cf. the clitic left dislocation structure in (5a) and the modification of the direct object in (5b).

- (5) a. #El cul_i ens el_i vam petar mirant la sèrie *Hotel Fawltly*.
 the butt us._{dat} it._{acc} past.1pl explode watching the series Hotel Fawltly
 ‘We laughed our butts off watching *Fawltly Towers*.’
 b. #Ens vam petar l’enorme cul que teníem mirant la sèrie
Hotel Fawltly.
 us._{dat} past.1pl explode the.big butt that had.1pl watching the series
 Hotel Fawltly
 ‘We laughed our butts off watching *Fawltly Towers*.’

The conceptual/metaphorical analysis of these idioms, which involves understanding/structuring an extreme intensity (target domain) in terms of an unreal, excessive detachment of a body part (source domain), is particularly interesting since it can be said to shed light on some important foundational tenets of Cognitive Linguistics: (i)

the existence of so-called “idiom families” is expected given the metaphorical nature of human thought (see Lakoff 1993). As shown above, conceptual metaphors can partly motivate *and* relate their meanings. As a result, these idioms are not to be regarded as mere isolated complex units having arbitrary meanings that turn out to be stored in our mental lexicon. (ii) The relevance and recurrence of body parts in many idiomatic constructions can be claimed to reflect the embodiment nature of human thought (see Johnson 1987; Gibbs 2005). (iii) Idioms are not to be regarded as peripheral products of language but can be shown to be structured according to core typological patterns: e.g., cf. the so-called “verb-framed pattern” of Romance languages (e.g., *treure el fetge per la boca (de tant treballar)*), which involves conflating directionality into the motion verb (*treure* ‘to get out’) and encoding manner as adjunct (e.g., *de tant treballar*), with the so-called “satellite-framed pattern” of Germanic languages (e.g., Engl. *to work one’s guts out*), which involves encoding the manner component into the verb and leaving the path/directionality component around the verb (as a “satellite”: e.g., the particle *out*). For further remarks concerning the application of Talmy’s (2000) famous bipartite typology of motion expressions (verb- vs. satellite-framed languages) to the realm of idioms, see Mateu/Esplan (2007).

To conclude this section, it will be important to make some relevant remarks on what a(n) (im)possible verbal idiom is, a question that is not usually addressed in merely taxonomic works. Indeed, there is already variety of interesting works in which syntactic and semantic constraints have been put forward: e.g., cf. O’Grady (1998) and Bruening (2017) for syntactic constraints, and Nunberg/Sag/Wasow (1994) for semantic ones. For example, consider the following typical constraint mentioned in the literature, which, due to its major theoretical relevance, has even been transformed into a hypothesis, i.e., the so-called “No Agent Idioms” hypothesis (e.g., see Harley/Stone 2013). It is often noted that there are few idioms, if any, that can be claimed to include agents. However, some examples that appear to contradict this can be found, such as the ones given in (6):

- (6) a. Un ocellet m’ ha dit que...
 a birdie me._{dat} has told that
 ‘A little bird told me that...’
 b. Se li ha menjat la llengua el gat.
 refl. you._{dat} has eaten the tongue the cat
 ‘The cat got her tongue.’

In addition, examples like (6a) have been said to be exceptional in the light of the relevant following constraint: i.e., there are many verb-object idioms, with an open slot for the subject (e.g., *fotre el camp*; cf. Engl. *to get the hell out*), but there appear to be very few, if any, subject-verb idioms with an open slot for the object (see Bruening 2017 for a syntactic explanation of this intriguing constraint). This author has in fact argued that examples like (6a) are not really problematic for this syntactic general-

ization since the nominal phrase *A little bird* can occur with this meaning with other verbs or even no verb (e.g., cf. *Un ocellet m'ha {explicat/comentat/xivat} que...* 'A little bird {explained/commented/squealed} to me that...'). Moreover, it is clear that the subject of the example in (6a) must have referential status (e.g., it can be pronominalized as ... *i jo li he respost que...* 'and I have replied to him/her that...'), whereby it cannot be interpreted as a property in Espinal's (2001) sense, unlike the direct object of true idioms like *fotre el camp*, where this nominal phrase does lack referentiality.

Another interesting case for testing the previous syntactic generalization could be presented by idioms like *riure-li els ulls* lit. 'her/his eyes smile', i.e., 'to be very happy'. Notice however that this apparently problematic example involves a case of inalienable possession and the nominal phrase *els ulls* 'the eyes' is probably not an external subject argument but an internal one bound by the dative possessive pronoun *li* 'to him/her', which would be higher in the relevant thematic hierarchy. Alternatively, following Nunberg/Sag/Wasow's (1994) semantic proposal, the fact that agents and dative possessors are typically not part of idioms is probably due to their particularly salient animate status. According to them, it is typical that "the nominal phrases of idiomatic phrases, on their *idiomatic* interpretation, tend not to have animate –or, more specifically, human – references" (1994, 528; emphasis theirs). But see O'Grady (1998, 305–309) for a critique of their semantic proposal.

Another typical constraint is that there are no idioms containing three clauses (perhaps for processing reasons). At most, one subordinate clause can be included: e.g., *no tenir on caure mort* lit. 'to have no place to fall dead', i.e., 'to be very poor' (cf. Engl. [*to have no pot [to piss in]*] 'to lack any standing').

Finally, it should also be noted that both idioms and idiomatic collocations can include syntactic adjuncts: cf. the PP *per la boca* lit. 'through the mouth' in *treure el fetge per la boca* (see above) and the PP *pels descosits* in *{parlar/dir-les/clavar-les} pels descosits* 'to be very talkative', respectively.

So far, we have dealt with verbal lexicalized phrases, which are the typical ones that are studied in classical works on idioms (e.g., see Nunberg/Sag/Wasow 1994). For reasons of space, the two following sections can only present a very brief sketch of nominal and other types of lexicalized phrases.

2.2 Lexicalized nominal phrases

As is often noted, a clear distinction between compounds and lexicalized nominal phrases is not always easy to draw. Traditionally, there are two different types of compounds: so-called "lexical compounds" (also known as "proper compounds", where only one word is involved: e.g., *rentaplats* 'dishwasher', *setciències* lit. seven-sciences, i.e., 'a person who thinks that s/he knows much more than other people', *pellroja* lit. skin-red, i.e., 'a person having a red skin', etc.) and "syntagmatic compounds" (also known as "improper compounds", when two words are involved: e.g.,

cotxe bomba ‘car bomb’, *camió cisterna*, ‘tank truck’, *home granota* ‘frog man’, etc.). For example, when dealing with N+N compounds, a distinction can be drawn on the basis of the placement of the plural inflection: it is attached to the end of the exocentric compound in the former type (e.g., *els pellroges*) and to the first word of the endocentric compound in the latter type (e.g., *els cotxes bomba*).

When dealing with so-called “root compounds” (i.e., complex words that are composed of two bare roots), it is also interesting to compare Catalan and English: for example, consider Cat. *home granota* and Engl. *frog man*. In Catalan this compound is lexically associated with the idiomatic meaning of an underwater diver. The typical case in Romance languages is that these root compounds are lexicalized nominal phrases. However, this is not the case in English; in this Germanic language this compound also has this lexicalized meaning but, unlike in Catalan, it can easily be used with many more meanings given the appropriate context: e.g., “<frog man> could mean the kind of man who looks like a frog, eats frogs, studies frogs, sells frogs, or wishes to be buried in a frog-shaped casket, for example” (Snyder 2016, 110). See also Padrosa-Trias (2010), for more discussion on this interesting typological phenomenon which is referred to in the literature as ‘The Compounding Parameter’.

In this chapter on lexicalized syntax, I will not deal with the distinction between different types of syntagmatic compounds (§7.2 Word-Formation), but rather will focus on showing how the diverse range of lexicalized nominal phrases can be profitably studied by taking the abovementioned distinction between idioms vs. (idiomatic and non-idiomatic) collocations into account. Recall that the key semantic property that distinguishes idioms from collocations is the absence or presence of compositionality, respectively. Consider, for example, an idiom like *ull de poll* ‘corn’ (foot condition), an idiomatic collocation like *tauleta de nit* ‘bedside table’, and a non-idiomatic collocation like *vaga de fam* ‘hunger strike’. The meaning of the first example is clearly non-compositional in the sense that it is not distributed among the parts of the idiom, i.e., the expression as a whole is mapped onto the meaning of the idiom. Furthermore, the two nominal elements (*ull* and *poll*) are idiomatic, whereby the conventionality relation involved is bidirectional (see above). In contrast, the meaning of the second and third examples of lexicalized nominal phrases can be claimed to be compositional since the meaning of these two collocations is distributed among the parts of the expression (*tauleta* and *vaga*, on the one hand, and *de nit* and *de fam*, on the other). Basically, the difference between the idiomatic collocation *tauleta de nit* and the non-idiomatic one *vaga de fam* has to do with the fact that *de nit*, but not *de fam*, is provided with considerable idiomatic burden. Notice also that the complex meaning of the lexicalized phrase *tauleta de nit* is not the mere result of combining/relating the meaning of ‘little table’ with that of ‘of night’ but a major denotative property is also added (cf. its lexical entry in the DIEC). Incidentally, the pair formed by Cat. *tauleta de nit* and Engl. *bedside table* provides a nice case to illustrate Langacker’s (2008) famous insight that meaning is a function of both conceptual content and semantic construal. According to this cognitive linguist, both expressions could be said to share the same

conceptual content (i.e., we are referring to the very same object) but differ in how they are semantically construed in each language (cf. the different semantic contribution of *de nit* vs. *bedside* in the two idiomatic collocations).

2.3 Other types of lexicalized elements

In this section, I briefly deal with other types of lexicalized phrases including prepositional, adverbial, and adjectival ones and point out that the abovementioned important distinction between idioms and (idiomatic and non-idiomatic) collocations can also be said to apply to this domain. For example, lexicalized adverbial phrases such as *en públic* ‘in public’ (cf. *parlar en públic* ‘speak in public’), *sens dubte* ‘without a doubt’ or adjectival ones like *sa i estalvi* ‘safe and sound’ can be analyzed as non-idiomatic collocations: their meaning is not mostly figurative and is clearly distributed among their parts. Prepositional phrases like *amb tots els ets i uts* or *amb tots els pèls i senyals* ‘with all the details’ are provided with a major idiomatic burden. This notwithstanding, these idiomatic expressions can also be analyzed as collocations to the extent that their meanings can also be distributed among their parts. For example, in these cases it is possible to distinguish the contribution of the literal meaning of the preposition *amb* ‘with’ from the idiomatic meanings contributed by the nominal expressions. In contrast, drawing such a distribution of meaning appears to be more difficult when dealing with lexicalized adverbial phrases like *de sobte* ‘all of a sudden’ (cf. *de cop i volta*), where the meaning can be said to be holistically associated with the entire expression. In this case, it may be advisable to treat them as prototypical idioms (i.e., as “idiomatic phrases” in the terminology of Nunberg/Sag/Wasow 1994).

The relevant distinction between idioms and idiomatic collocations can also be exemplified with the complex prepositional unit *en nom de* lit. ‘in the name of’: cf. its compositional meaning in *En Joan vindrà en nom de la Maria* lit. ‘Joan will come in the name of Maria, i.e., on behalf of Maria’ and its less compositional meaning in *En nom del Pare, del Fill i de l'Esperit Sant* ‘In the name of the Father, the Son, and the Holy Spirit’. The first use can be claimed to correspond to that of an idiomatic collocation since the noun *nom* ‘name’ acquires the idiomatic sense of *representació* only in this context: *en nom de* – *en representació de* ‘on behalf of’. The second use most probably corresponds to that of an idiom to the extent that the expression of *En nom de* can be claimed to be holistically mapped onto the idiomatic meaning of invocation, which is in turn associated with the making the sign of the cross.

Finally, it is also worth noting that the existence of complex conjunctions like *atès que* ‘given that, since’, *{llevat/tret/excepte} que* ‘except that’ or *de manera que* ‘in such a way that’, among others, is often attributed to the phenomenon of grammaticalization, i.e., lexical expressions becoming part of grammatical elements like complex conjunctions. This phenomenon is then different from that of lexicalization, i.e., syntactic expressions becoming part of lexical units (see above). In any case, it is clear

that these complex conjunctions are also stored as units in our mental lexicon (see Jackendoff 1997).

3 Concluding remarks

In this chapter I have concentrated on the distinction between idioms and collocations and have applied it to some basic types of lexicalized phrases in Catalan, with a main focus on verbal phrases. Following Nunberg/Sag/Wasow (1994), three different semantic properties must be distinguished when dealing with the interpretation of lexicalized phrases: conventionality, compositionality, and transparency. Strictly speaking, only those lexicalized phrases whose meaning is non-compositional can be classified as idioms (also known as “idiomatic phrases” in Nunberg/Sag/Wasow 1994). In contrast, those expressions that are idiomatic but do have a compositional meaning are better referred to as idiomatic collocations (cf. “idiomatically combining expressions” in Nunberg/Sag/Wasow 1994). The semantic property of (non-)compositionality has been shown to be correlated with the syntactic property of (in-)flexibility: for example, the syntax of idioms like *trencar-se el cap* ‘to break one’s head’, whose meaning is non-compositional (i.e., their meaning is not distributed among the parts of the expression but can be said to be holistically associated with it), has been shown to be far less flexible than the syntax of idiomatic collocations like *trencar el gel* ‘to break the ice’. Furthermore, the conceptual property of transparency/opacity has been argued to be related to the presence/absence of metaphorical motivation (see Lakoff 1993; Gibbs 1995). For example, as shown by Espinal/Mateu (2010), idioms like *treure el fetge per la boca* ‘to expel the liver through the mouth’ are semantically non-compositional but can be claimed to be conceptually transparent by positing the existence of a relevant conceptual metaphor that relates a physical source domain (an excessive, unreal extraction or detachment of a body part) to an abstract target domain (an extreme intensity). As emphasized by Lakoff and Gibbs, the existence of conceptual metaphors explains why the meanings of many idioms are much more complex than their associated definitions that are typically found in dictionaries (cf. Raspall/Martí 1984; Espinal 2004; 2005). Indeed, the meaning of *treure el fetge per la boca* is not just “to be(come) very tired” (nor the meaning of Engl. *to cry one’s eyes out* is just ‘to cry a lot’, Jackendoff 1997).

Finally, it should be noted that I have dealt with those phraseological units that, in Corpas’ (1996) classification, do not involve an utterance. According to this author, the set of so-called “phraseological utterances” can be claimed to be divided into two basic types: proverbs (*Qui la fa, la paga* ‘whoever does it, pays for it’) and routine formulae (*Per molts anys!* Cf. Engl. *Many happy returns!*). As is well-known, proverbs are the object of investigation of paremiology (for example, see Conca 1987). Proverbs are general statements that are believed to express a universal truth and are most significantly connected with aspects of culture-based social interaction. By using

proverbs, one refers to socially approved ideas that can be used instead of a more formal argumentation. Routine formulae are also referred to in the literature as “communicative phrasemes” or “pragmatic idioms”. Indeed, it should be obvious why routine formulae are restricted to the pragmatic level; they are tools of communication and their most important function is also the constitution of speech acts (typically, greetings, expressions of thanks, excuses, etc. belong to the core elements of this class). They are therefore part of a larger complex of stereotyped action patterns and social interactions (for further relevant discussion, see Piirainen 2008).

4 Bibliography

- Bosque, Ignacio (2001), *Sobre el concepto de “colocación” y sus límites*, *Lingüística Española Actual* 23/1, 9–40.
- Bruening, Benjamin (2017), *Syntactic Constraints on Idioms (Do Not Include Locality)*, in: Claire Halpert et al. (edd.), *A Pesky Set: Papers for David Pesetsky*, *MIT Working Papers in Linguistics*, Cambridge MA, MIT Press, 183–192.
- Conca, Maria (1987), *Paremiologia*, València, Universitat de València.
- Corpas, Gloria (1996), *Manual de fraseología española*, Madrid, Gredos.
- Croft, William/Cruse, D. Alan (2004), *From Idioms to Construction Grammar*, in: William Croft/D. Alan Cruse, *Cognitive Linguistics*, Cambridge, Cambridge University Press, 225–256.
- DIEC (2007) = Institut d'Estudis Catalans, *Diccionari de la llengua catalana*, Barcelona, Edicions 62/ Enciclopèdia Catalana.
- Espinal, M. Teresa (2001), *Property Denoting Objects in Idiomatic Constructions*, in: Yves D'Hulst et al. (edd.), *Romance Languages and Linguistic Theory 1999. Selected Papers from Going Romance*, Amsterdam/Philadelphia, Benjamins, 117–141.
- Espinal, M. Teresa (2002), *Idiomatic Constructions vs. Light Verb Constructions*, in: Manuel Leonetti et al. (edd.), *Current Issues in Generative Grammar. 10th Colloquium on Generative Grammar Selected Papers*, Alcalá de Henares, Universidad de Alcalá, 69–81.
- Espinal, M. Teresa (2004), *Diccionari de sinònims de frases fetes*, Barcelona/València, Universitat Autònoma de Barcelona/Publicacions de l'Abadia de Montserrat/ Universitat de València, <http://ddd.uab.cat/record/89642> (last accessed: 30.11.2018).
- Espinal, M. Teresa (2005), *A Conceptual Dictionary of Catalan Idioms*, *International Journal of Lexicography* 18/4, 509–540. DOI: 10.1093/ijl/eci034.
- Espinal, M. Teresa/Mateu, Jaume (2005), *La llengua a les frases fetes*, *Articles de Didàctica de la Llengua i de la Literatura* 36 (special issue: *Fraseologia i educació discursiva*), 20–31.
- Espinal, M. Teresa /Mateu, Jaume (2010), *On Classes of Idioms and Their Interpretation*, *Journal of Pragmatics* 42, 1397–1411. DOI: 10.1016/j.pragma.2009.09.016.
- Espinal, M. Teresa/Mateu, Jaume (2011), *Bare Nominals and Argument Structure in Catalan and Spanish*, *The Linguistic Review* 28, 1–39. DOI: 10.1515/tlir.2011.001.
- Gibbs, Raymond W. (1995), *Idiomaticity and Human Cognition*, in: Martin Everaert et al. (edd.), *Idioms: Structural and Psychological Perspectives*, Hillsdale, Lawrence Erlbaum Associates, 97–116.
- Gibbs, Raymond W. (2005), *Embodiment and Cognitive Science*, Cambridge, Cambridge University Press.
- Ginebra, Jordi (2000), *Sintaxi i fraseologia: els límits de les unitats fraseològiques verbals*, in: Vicent Salvador/Adolf Piquer (edd.), *El discurs prefabricat: Estudis de fraseologia teòrica i aplicada*, Castelló de la Plana, Universitat Jaume I, 65–80.

- Ginebra, Jordi (2003), *Fraseologia, concurrències lèxiques i llengua estàndard*, in: Miquel Àngel Pradilla (ed.), *Identitat lingüística i estandardització*, Valls, Cossetània, 7–55.
- Ginebra, Jordi (2008), *Els verbs de suport en català i en anglès: estudi contrastiu a partir d'un petit corpus paral·lel*, *Els Marges* 85, 53–72.
- Ginebra, Jordi (2017), *Lexical Combinatorics in Catalan*, in: Sergi Torner/Elisenda Bernal (edd.), *Collocations and Other Lexical Combinations in Spanish. Theoretical, Lexicographical and Applied Perspectives*, London, Routledge, 305–314.
- Ginebra, Jordi/Navarro, Pere (2015), *Concurrències lèxiques en català i en espanyol: uns quants contrastos*, in: Àlex Martín Escribà/Adolf Piquer Vidal/Fernando Sánchez Miret (edd.), *Actes del Setzè Col·loqui Internacional de Llengua i Literatura Catalanes. Universitat de Salamanca, 1–6 de juliol de 2012*, vol. 2, Barcelona, Publicacions de l'Abadia de Montserrat, 217–228.
- Glasbey, Sheila R. (2007), *Aspectual Composition in Idioms*, in: Louis de Saussure/Jacques Moeschler/Genoveva Puskás (edd.), *Recent Advances in the Syntax and Semantics of Tense, Aspect and Modality*, Berlin/New York, Mouton de Gruyter, 71–88. DOI: 10.1515/9783110198768.71.
- Hale, Kenneth L./Keyser, Samuel J. (1993), *On Argument Structure and the Lexical Expression of Syntactic Relations*, in: Kenneth L. Hale/Samuel J. Keyser (edd.), *The View from Building 20: Essays in Linguistics in Honor of Sylvain Bromberger*, Cambridge, MA, MIT Press, 53–109.
- Hale, Kenneth L./Keyser, Samuel J. (2002), *Prolegomenon to a Theory of Argument Structure*, Cambridge, MA, MIT Press.
- Harley, Heidi/Stone, Megan (2013), *The “No Agent Idioms” Hypothesis*, in: Raffaella Folli/Christina Sevdali/Robert Truswell (edd.), *Syntax and its Limits*, Oxford, Oxford University Press, 283–311.
- Hausmann, Franz Josef (1997), *Tout est idiomatique dans les langues*, in: Michel Martins-Baltar (ed.), *La locution entre langue et usages*, Paris, ENS Éditions Fontenay/St. Cloud, 277–290.
- Jackendoff, Ray (1997), *Idioms and Other Fixed Expressions*, in: Ray Jackendoff, *The Architecture of the Language Faculty*, Cambridge, MA, MIT Press, 153–178.
- Johnson, Mark (1987), *The Body in the Mind. The Bodily Basis of Meaning, Imagination, and Meaning*, Chicago, The University of Chicago Press.
- Kiparsky, Paul (1997), *Remarks on Denominal Verbs*, in: Àlex Alsina et al. (edd.), *Complex Predicates*, Stanford, CSLI Publications, 473–499.
- Lakoff, George (1993), *The Contemporary Theory of Metaphor*, in: Andrew Ortony (ed.), *Metaphor and Thought*, Cambridge, Cambridge University Press, 202–251.
- Langacker, Ronald W. (2008), *Cognitive Grammar: A Basic Introduction*, Oxford/New York, Oxford University Press.
- Lorente, Mercè (2002), *Altres elements lèxics*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 1, Barcelona, Empúries, 831–883.
- Mateu, Jaume/Espinal, M. Teresa (2007), *Argument Structure and Compositionality in Idiomatic Constructions*, *The Linguistic Review* 24, 33–59. DOI: 10.1515/TLR.2007.002.
- Mateu, Jaume/Rigau, Gemma (2010), *Verb-Particle Constructions in Romance: A Lexical-Syntactic Account*, *Probus* 22, 241–269. DOI:10.1515/prbs.2010.009.
- Mel'čuk, Igor (1995), *Phrasemes in Language and Phraseology in Linguistics*, in: Martin Everaert et al. (edd.), *Idioms: Structural and Psychological Perspectives*, Hillsdale, Lawrence Erlbaum Associates, 167–232.
- Mendivil, José Luis (1999), *Las palabras disgregadas. Sintaxis de las expresiones idiomáticas y los predicados complejos*, Zaragoza, Prensas Universitarias de Zaragoza.
- Mendivil, José Luis (2009), *Palabras con estructura externa*, in: Elena de Miguel (ed.), *Panorama de la lexicología*, Barcelona, Ariel, 83–113.
- Mestres, Josep M. (2007), *Per un tractament unívoc de les unitats pluriverbals en fraseologia i terminologia*, in: Mercè Lorente et al. (edd.), *Estudis de lingüística i de lingüística aplicada en*

- honor de M. Teresa Cabré Castellví*, Barcelona, Institut Universitari de Lingüística Aplicada, 381–400.
- Nunberg, Geoffrey/Sag, Ivan/Wasow, Thomas (1994), *Idioms*, *Language* 70, 491–538.
- O'Grady, William (1998), *The Syntax of Idioms*, *Natural Language and Linguistic Theory* 16, 279–312. DOI:10.1023/A:1005932710202.
- Padrosa-Trias, Susanna (2010), *Complex Word-Formation and the Morphology-Syntax Interface*, unpublished doctoral dissertation, Bellaterra, Universitat Autònoma de Barcelona.
- Piirainen, Elisabeth (2008), *Figurative Phraseology and Culture*, in: Sylviane Granger/Fanny Meunier (edd.), *Phraseology. An Interdisciplinary Perspective*, Amsterdam/Philadelphia, Benjamins, 207–228.
- Raspall, Joana/Martí, Joan (²1984), *Diccionari de locucions i frases fetes*, Barcelona, Edicions 62.
- Ruwet, Nicholas (1991), *Syntax and Human Experience*, edited and translated into English by John Goldsmith, Chicago/London, The University of Chicago Press.
- Salvador, Vicent (1995), *De la fraseologia a la lingüística aplicada*, *Caplletra* 18, 11–30.
- Snyder, William (2016), *Compound Word Formation*, in: Jeffrey Lidz et al. (edd.), *The Oxford Handbook of Developmental Linguistics*, Oxford/New York, Oxford University Press, 89–110.
- Talmy, Leonard (2000), *Toward a Cognitive Semantics*, Cambridge, MA, MIT Press.
- Titone, Debra A./Connine, Cynthia M. (1999), *On the Compositional and Noncompositional Nature of Idiomatic Expressions*, *Journal of Pragmatics* 3/12, 1655–1674. DOI:10.1016/S0378-2166(99)00008-9.

Maria Josep Cuenca

6 Pragmatics and Text Linguistics

Abstract: This chapter describes some relevant aspects of the pragmatics and text linguistics of Catalan, focusing on the interaction between appropriateness and grammar, both at the sentence and the text level. Firstly, the use of the definite article with personal names illustrates the importance of shared knowledge and presuppositions in discourse construction. Secondly, I describe the form and use of Catalan demonstratives and analyse their use and interaction with other reference devices by identifying the translation strategies used to convey English demonstratives in Catalan. Subjectivity and modalisation are another area of interest for pragmatics and text linguistics in which modal particles play an important role. This is illustrated by describing the use of the modal particle *pas* and its discourse effects. Finally, the chapter examines discourse markers, specifically parenthetical markers and pragmatics markers as connective devices, and summarises the uses of the marker *home* ‘man’.

Keywords: pragmatics, text linguistics, modal particle *pas*, demonstratives, names, article, discourse markers

1 Pragmatics and Text Linguistics of Catalan

Pragmatics and text linguistics are two interrelated areas of analysis that share a language-in-use approach. They have attracted growing interest in Catalan linguistics since the 1980s, by which time the use of Catalan had entered a new stage in which teaching the language both at schools and universities became general after the dark years of the Franco dictatorship. A steady increase in the number of publications in these areas during the 1990s led to consolidation in the 21st century, as the books *Les claus de la pragmàtica* (Bassols 2001) and *The Pragmatics of Catalan* (Payrató/Cots 2011) show.¹

¹ There are several panoramic overviews of this area of study: Boix/Payrató (1995, 2007) review the main contributions to Catalan sociolinguistics and pragmatics in the period 1989–1996, while Payrató (2016) continues the state of the art from 1997 to 2012. Alturo (2003; 2011) presents a synthetic history of pragmatic studies on Catalan, including sociolinguistics (2003) and discourse analysis research (2011), and Bassols (2003) focuses on pragmatics and discourse analysis. The collective volume edited by Payrató/Cots (2011) is an overview of research in Catalan pragmatics, which includes papers on the interaction between grammar and use, as well as on specific text and discourse aspects, along with some sociolinguistic and variationist contributions.

From a general perspective, a text can be understood as a communicative unit that can be described by taking into consideration several key properties (cf., e.g., Halliday/Hasan 1976; de Beaugrande/Dressler 1981). The three main properties that allow us to analyse a text are appropriateness, coherence and cohesion (cf. Castellà 1992; Cuenca 2008a). Appropriateness refers to the relationship between the text as a linguistic unit and the extralinguistic factors that influence its production and interpretation, such as the time or place of production, language variety, register, interlocutors and other discourse voices and subjectivity. Coherence results from the interaction between textual meanings and how they are organised. Cohesion is syntax at the level of the text and includes the mechanisms that tie together different parts of text or relate text and context, namely, reference (i.e. deictic and phoric processes), modality and connection. These mechanisms are denoted by lexical and grammatical markers that explicitly indicate appropriateness and coherence phenomena.

It would be impossible to account for all the pragmatic and text level processes in Catalan in just one chapter. In addition, some processes are common to other languages so there is little point in describing them as if they were specific to Catalan. In this chapter, I will focus on the interaction between appropriateness and grammar at both the sentence and text level (i.e. cohesion). In the following sections, I will briefly review and exemplify some outstanding phenomena with features specific to Catalan, which fall into the following four areas, namely, (i) pragmatic meaning and its relationship with grammar, (ii) deixis from a discourse perspective, (iii) subjectivity and modality, and (iv) discourse marking.

2 Shared knowledge and presuppositions: The use of the article with personal names

Pragmatic meaning is the result of the interaction between semantic or propositional meaning and the meaning effects related to the communicative situation, especially to what the addressor and the addressee know, presuppose and assume as common or new knowledge. This kind of meaning is often implicit, but is sometimes tied to the use of grammatical or lexical markers.

In the case of Catalan, definite articles can express pragmatic meaning when used with personal names. Proper nouns corresponding to first names, last names or full names can be preceded by the general definite article (*el, la* ‘the’ inflected for male and female names, respectively), derived from the Latin distal demonstrative *ILLE* (‘that’), or by the so-called personal article (*en, na*), which can only specify anthroponyms and derives from the Latin noun *DOMINE-DOMINA* ‘sir/madam’ (GIEC 2016, 16.3.1; Brucart 2002, 7.3.4; 75.2 The Simple Sentence).

The distribution and use of the article preceding proper nouns are not the same in all Catalan-speaking territories: in Catalonia varieties, male names are generally

preceded by *en* or *el* (which is the only option if the name begins with a vowel), whereas female ones are preceded by *la* (1). In the Balearic Islands varieties, the personal article usually precedes both male and female names (2).²

- (1) He convidat a la festa *el* Joan, *l'Àlex*, *en* Carles i *la* Maria
'I have invited (the) John, (the) Alex, (the) Charles and (the) Maria to the party'
- (2) He convidat a la festa *en* Joan, *n'Àlex*, *en* Carles i *na* Maria
'I have invited (the) John, (the) Alex, (the) Charles and (the) Maria to the party'

However, when referring to public people (such as authors, politicians, famous people), to foreigners or to characters from the Bible or mythology, the default option is to avoid the article:

- (3) a. Enric Granados fou compositor i pianista
'Enric Granados was a composer and piano player'
- b. Aristòtil és un gran filòsof
'Aristotle is a great philosopher'
- c. Marie Curie es va llicenciar en física l'any 1893
'Marie Curie got her degree in physics in 1893'
- d. Samsó fou el darrer dels set jutges d'Israel
'Samson was the last of the seven judges of Israel'

The pragmatic difference introduced by the article can be seen by comparing the examples in (4), where the former (4a) refers to someone that both speaker and hearer are familiar with and the latter (4b) indicates that the person (*Maria Puig*) is considered to exist outside the interlocutors' private sphere and is (most likely) not an acquaintance of theirs.

- (4) a. La/Na Maria farà el discurs i agafarà un avió per a tornar a casa tot seguit
'(The) Maria will deliver the address and fly back home immediately'
- b. Maria Puig farà el discurs inaugural
'Maria Puig will deliver the inaugural address'

The different presuppositions related to the presence or absence of the article explain why the use of the article is very frequent in narrative texts and is infrequent in

² In the Southern dialects of Catalonia, namely in Tortosa's variety, as well as in Valencian dialects it is not common to use an article before anthroponyms.

newspapers, for instance, where there is a strong tendency to avoid the article preceding a name (whether full, last or first), as the following newspaper headlines (from newspaper *Ara*) show:

- (5) a. Neus Munté, Carlota Pi, Estel Solé, Ada Parellada, Teresa Estrach, Carme Lloveras i Esther Vera expliquen el que els passa en el dia a dia
‘Neus Munté, Carlota Pi, Estel Solé, Ada Parellada, Teresa Estrach, Carme Lloveras and Esther Vera explain what happens to them in their everyday life’
- b. Millet es planteja confessar però el fiscal en té prou amb Montull
‘Millet is considering confessing but the prosecutor is satisfied with Montull’
- c. Qui pot i qui no podrà substituir Luis Enrique?
‘Who can and who could not replace Luis Enrique?’

It can be concluded that the presence or absence of the article with names is pragmatically conditioned, reflecting familiarity with the person being referred to. The presence of the article indicates that the third person belongs to the speaker’s private sphere, and that the hearer is supposed to identify this person either contextually or because of shared knowledge (Wheeler/Yates/Dols 1999, 3.3.2).

As a result, the use of the article in contexts that typically include omission activates the inference that either the person named is considered to be familiar to the speaker or that there is some kind of emotional involvement, whether admirable (6a) or not (6b). This is often the case with well-known Catalan figures or with present-day famous figures of other origins:

- (6) a. La producció literària de *la* Rodoreda és impressionant
‘(The) Rodoreda’s literary production is impressive’
- b. *En* Trump n’ha dit una altra de ben increïble
‘(The) Trump has said something incredible again’

Yet, it would be awkward to use the article with reference to a historical figure (7a), unless the proper noun is used as a common noun (7b).

- (7) a. *!El* Beethoven no m’agrada
‘I don’t like the Beethoven’
- b. *El* darrer Beethoven és molt més romàntic que el primer
‘The later Beethoven is much more romantic than his early work’

In brief, the use of articles preceding names in Catalan allows the speaker to introduce pragmatic information related to presuppositions and shared-knowledge. Only pragmatics can explain this use of the article, which from a purely syntactic perspective

would be exceptional, since proper nouns are definite per se when used in discourse, thus making a determiner syntactically redundant.³

3 Deixis: The form and use of Catalan demonstratives

Deixis is a cohesive device by which linguistic items are used to point to the situational context of communication. There are personal deictics, such as personal pronouns and possessives corresponding to or including the addressor (*jo* ‘I’, *nosaltres* ‘we’, *nostre* ‘our’) or the addressee (*tu* ‘you-singular’, *vosaltres* ‘you-plural’, *teu/vostre* ‘yours’), temporal deictics, including certain adverbs (e.g., *ara* ‘now’, *demà* ‘tomorrow’), and space deictics which also include certain adverbs (*aquí* ‘here’, *allí* ‘there’), as well as noun specifiers and pronouns, which fall under the general concept of demonstratives (see Payrató 2002; Nogué 2008; 2011; 2015; Cuenca 2008a, chap. II.1; and also 75.2 The Simple Sentence). Demonstratives work jointly with phoric markers, such as third person pronouns, and with ellipsis to maintain topic reference in discourse (on phoric marking in Catalan see, e.g., Rigau 1981; Cuenca 2008a, chap. III; and Ribera 2012). For instance, in (8) the topic *llibre* is introduced with the demonstrative *aquest* ‘this’, then retrieved by a phoric weak pronoun, *e(l)* ‘it’ and is finally deleted in subject position.

- (8) Agafa *aquest* llibre i porta’l a la biblioteca. És prestat.
 ‘Take this book and take it to the library. (It) is borrowed’

Demonstratives are deictic units that typically point to elements of the situational context of an utterance, thus marking spatial distance with regard to the deictic origin, which generally coincides with the addressor’s position. Demonstratives in Catalan follow specific patterns of behaviour, which are different in various Catalan dialects. In the following, the space deixis systems in Catalan (i.e., three- and two-term) will be briefly described. The use of the two-term system will be illustrated by looking at the translation of English demonstratives into Catalan.

³ It should be noted that, although grammars coincide in their identification of the alternation described in this section (e.g. GIEC 2016, 16.3.1.2; Brucart 2002, 7.3.4), corpus analysis reveals a certain degree of variation. In dialects where the use of the article is pervasive in informal speech, some speakers tend to extend this use to all kinds of nouns and even in Spanish, although in Spanish the use of the article with names is considered too informal and mostly incorrect. More research is needed to establish what grammars describe and prescribe in relation to actual use.

Catalan exhibits two space deixis systems: a three-term or ternary system and a two-term or binary one. The three-term system is the classical system and is still in use in Valencian dialects and other smaller regions, such as Eivissa and la Franja of Aragon. The two-term system is used in the rest of the Catalan dialects (see Badia i Margarit 1994, 497–502; Payrató 2002, 1165–1169; Nogué 2015, 1; and GIEC 2016, 16.4.).

It should be pointed out that the Catalan binary system does not coincide with the English binary system, which clearly distinguishes between different spaces, as synthesised in Table 1, where for simplicity only demonstrative determiners are included.

Table 1: Deictic demonstrative systems in Catalan and English.

Deictic space		Proximal		Distal
		Speaker (‘this’)	Hearer (‘that-near’)	Others (‘that-far’)
Catalan	three-term system	<i>este/aquest</i>	<i>eixe/aqueix</i>	<i>aquell</i>
	two-term system	<i>aquest</i>		<i>aquell</i>
English		<i>this, these</i>	<i>that, those</i>	

The three-term system includes three degrees of proximity to the speaker: immediate proximity (expressed by the simple form *este* and its inflected variants *esta*, *estos*, *estes*, or by the reinforced demonstratives *aquest*, *aquesta*, *aquests*, *aquestes*), intermediate proximity (*eixe/aqueix* and inflected variants), which often correspond to proximity to the hearer, and distance from both speaker and hearer (*aquell* and inflected variants). We can thus say:

- (9) a. Agafa *aquest* llibre
 ‘Take this book’ (that I am holding)
 b. Dona’m *aqueix* llibre
 ‘Give me that book’ (that you are holding)
 c. Porta *aquell* llibre
 ‘Bring that book’ (which is far away from both of us)

The Catalan two-term system clearly distinguishes the space shared by speaker and hearer, expressed by proximal demonstratives (*aquest*, *aquesta*, *aquests*, *aquestes*), to the non-shared space, expressed by distal demonstratives (*aquell*, *aquella*, *aquells*, *aquelles*). Therefore, the previous sentences would be expressed as follows:

- (10) a. Agafa *aquest* llibre
 ‘Take this book’ (that I am holding)
 b. Dona’m *aquest* llibre
 ‘Give me that book’ (that you are holding)
 c. Porta *aquell* llibre
 ‘Bring that book’ (which is far away from both of us)

Therefore, in the binary system, the proximal demonstrative, *aquest*, is tied to both speaker and hearer. The distal demonstrative, *aquell*, points to items that are distant from the interlocutors. This distribution clearly differs from the English binary system, which distinguishes between the speaker’s space and the non-speaker’s space, as the glosses in the previous examples show.

Demonstratives, however, do not only refer to extra-linguistic context. There are cases where demonstratives point to elements in the discourse context or co-text, mainly to antecedents previously introduced in the discourse. The difference between situational and non-situational demonstratives can be seen in (11).

- (11) a. He comprat *aquest* llibre
 ‘I bought this book’ (*this book* = a book that is near me or that I am holding)
 b. Li compraré *la gramàtica de l’IEC*, perquè ell vol *aquest* llibre
 ‘I will buy him IEC grammar, because he wants this book’ (*this book* = *IEC grammar*)

In (11a), *aquest* points to an object in the situational context and acts as a situational demonstrative. In (11b), *aquest* does not point to an object in the situational context but to a noun phrase previously introduced in the discourse, *IEC grammar*, and does not indicate proximity in space but in the text.

The use of Catalan demonstratives in discourse can be better understood by comparing it with English. The analysis of the translation of demonstratives in English fiction into Catalan included in Ribera i Condomina/Cuenca (2013) highlights some interesting facts connecting the binary systems of both languages and the use of Catalan demonstratives in general. In the corpus analysed in Ribera i Condomina/Cuenca (2013),⁴ four translation strategies have been identified:

⁴ The corpus consists of two novels: *Through a Glass Darkly* by Donna Leon (London, Arrow Books, 2006 [Glass]) and *The Interpretation of Murder* by Jed Rubenfeld (London, Headline Review, 2006 [Murder]), and their Catalan translations (*Cristall enverinat*, Barcelona, Edicions 62, 2006 [Cristall])/ *La interpretació del crim*, Barcelona, Edicions 62, 2007 [Crim]).

- (a) Deictic maintenance implies literal translation, that is, a proximal demonstrative translated by a proximal demonstrative (PD > PD) and a distal one by a distal one (DD > DD).
- (b) Deictic shift occurs when a proximal demonstrative is translated by a distal one (PD > DD) or *vice versa* (DD > PD), so that the deictic point of reference changes.
- (c) Deictic neutralisation accounts for the cases in which a demonstrative in the source text (English) is deleted or translated by a non-deictic unit, usually a definite article or pronoun (D > 0).
- (d) In the Catalan target text, some demonstratives have been added to the English text. This is what Ribera i Condomina/Cuenca (2013) call *over-marking*.

Let us describe and illustrate the four strategies in turn. (The examples come from Ribera i Condomina/Cuenca 2013).

a) Deictic maintenance is relatively frequent with situational proximal demonstratives (58.3%). This strategy accounts for the translation of 31.7 % of all English demonstratives.

(12)

The coroner said he could by no means allow it: in cases of homicide, the decedent's body must by law be taken into custody for an autopsy. 'Not <i>this</i> body', answered Banwell (Murder, 27).	El forense li va dir que no ho podia permetre de cap manera: en casos d'homicidi, el cos del difunt quedava custodiat fins que li practicaven l'autòpsia. —Doncs <i>aquest</i> ['this'] cos no —va dir en Banwell (Crim, 30).

In the case of situational deictics, as in (12), proximal demonstratives both in English (*this body*) and Catalan (*aquest cos*) point to the addressor's deictic space and thus there is no change of deictic centre.

Deictic maintenance is also possible but less frequent when demonstratives are non-situational, as *these cabs* in (13) referred to the previous noun-phrase *gasoline cars for hire*.

(13)

In 1907, the New York Taxicab Company launched the first fleet of gasoline cars for hire, equipped with meters so that riders could see the fare. <i>These</i> cabs were instant hits (Murder, 224).	El 1907, la New York Taxicab Company va presentar la primera flota de cotxes amb gasolina i amb taxímetres incorporats perquè els usuaris poguessin veure la tarifa. <i>Aquests</i> ['these'] taxis van tenir un èxit immediat (Crim, 191).

b) Deictic shift is triggered by discourse factors and mainly affects distal deictics. This strategy accounts for 28.2 % of the English demonstratives. The shift from a proximal English demonstrative to a distal Catalan one (14) is not very frequent but is significant in fiction.

(14)

'I teach at the university,' she said. Paola had never mentioned anyone like <i>this</i> young woman, but that did not necessarily mean anything (Glass, 41).	—Sóc professora a la universitat. La Paola no li havia parlat mai de ningú de les característiques d' <i>aquella</i> ['that'] jove, però això no havia de voler dir res necessàriament (Cristall, 38)

In (14), the use of a proximal deictic in the narrator's discourse points to the hypothesis that English authors often use proximal deictics in narrative sequences to create identification with the main character. In Catalan, on the contrary, there is a strong tendency to shift the deictic centre and select a distal demonstrative, which erases the narrator's identification with the character. In other words, Catalan foregrounds the emotional and/or temporal distance with respect to the facts narrated, whereas English tends to highlight the character's point of view.

Shifts from a distal English demonstrative to a proximal Catalan one are especially frequent in situational uses (15) or when an addressor is referring back to his or her addressee's words (16).

(15)

... Usually they served the sort of still red that came in large bottles, and instead of the thin glass he had in his hand, the wine was served in plastic cups. 'And <i>those</i> ?' she asked (Glass, 40).	Normalment servien aquell vi negre sense efervescència que venia en ampolles grosses, i en lloc de la delicada copa que tenia a la mà el distribuïen en gots de plàstic. —¿I <i>aquestes</i> ['these'] peces? —va preguntar ella (Cristall, 37).

(16)

'Weak little idiot.'	—Idiota i bleida.
'You wouldn't say <i>that</i> if you talked to her,' he said (Murder, 101).	—No ho diria pas, <i>això</i> ['this'], si parlés amb ella (Crim, 92).

This tendency correlates with the different distribution of deictic space in Catalan, where proximal deictics also point to the hearer's space and thus correspond to English distal deictics pointing to the hearer's sphere.

c) Some English demonstratives cannot be translated with demonstratives in Catalan. The demonstrative is deleted or translated with a non-deictic unit, mainly a definite article or pronoun.

(17)

'I'll be honest with you, man: if I took us straight back down, right now, all the way down, I might just save you.' <i>This</i> was true (Murder, 182).	—Et seré sincer. Si tornéssim a baixar de seguida, ara mateix, fins a baix de tot, potser encara et salvaria. —[Ø] Era veritat (Crim, 157).

In (17) the demonstrative *this* is deleted (*this was true/era veritat*) because Catalan is a pro-drop language and, therefore, a subject pronoun is only possible in contrastive contexts, which is not the case here. Neutralisation accounts for 40.1 % of all English demonstratives, making it the most frequent translation strategy in the analysed corpus.

Neutralisation implies a loss of deictic force and sometimes also of the empathetic nuance of the English original, which affects the involvement of the character or the narrator in the narration. It is more frequent with distal deictics (58.1 % of the cases) and it is predominant with non-situational demonstratives (181 cases out of 196 neutralised demonstratives, 92.3 %).

d) Some demonstratives are added to the English original. This strategy of deictic overmarking accounts for 38.9% of all demonstratives identified in the Catalan translation, and it is more frequent when demonstratives are non-situational (91.3%) and with distal forms (63%).

(18)

<i>The</i> two figures represented, for Littlemore, a world to which he had no access (Murder, 137).	Per a en Littlemore, <i>aquelles</i> ['those'] dues figures representaven un món on ell no tenia accés (Crim, 121).

In (18), the definite article *the* has been translated with a distal demonstrative (*aquelles*), although a literal translation could have been possible. The example shows how overmarking adds discourse effects related to the speaker's, either the narrator's or a character's, perspective.

Overmarking often applies to an English phoric pronoun, especially *it*, in contexts where neither deletion nor pronominalisation is possible in the target language (19).

(19)

The man in black tie doubted this assertion very much, although it happened to be true. But true or not, <i>it</i> made no difference, because Malley had seen him now (Murder, 179).	L'home de la corbata negra no es va acabar de creure la resposta, tot i que va resultar que era veritat. Però ara <i>això</i> ['this'] tant era, perquè en Malley l'havia vist (Crim, 154).

Overmarking is also used to avoid referential ambiguity in the target language (20).

(20)

As he watched, this maestro became the maestro for whom his father had worked. And as Brunetti continued to watch, <i>he</i> became every maestro who had worked the glass for more than a thousand years (Glass, 14).	Mentre se'ls mirava, aquell maestro es convertí de sobte en el maestro per qui havia treballat el seu pare. Al cap d'una estona, <i>aquell</i> ['that'] home ja s'havia convertit en la imatge de tots els mestri que havien treballat el vidre des de feia més de mil anys (Cristall, 110).

In (20), the antecedent *this maestro* is retrieved in English by a subject pronoun (*he*). In the Catalan version, the subject pronoun *ell* would be ambiguous (“Al cap d’una estona, *ell* ja s’havia convertit...”), since it could refer to the *maestro* or to *Brunetti*. The solution adopted is a noun phrase including a demonstrative followed by a synonym or a hyperonym (*aquell home* ‘that man’).

To sum up, Catalan demonstratives exhibit complex behaviour both descriptively and cross-linguistically. The deictic system is a key factor in the translation of situational demonstratives, but it is not so important in the case of non-situational ones. Non-situational demonstratives alternate with other reference devices and are often substituted with prototypical phoric devices such as ellipsis and anaphora. The selection of a specific reference device depends on constraints “related to the degree of accessibility of the intended referent, including some syntactic factors (e.g. referents coded as subjects are more salient and thus more frequently deleted in pro-drop languages), the cognitive status of the referents and their degree of activation at a given point in the discourse” (Ribera i Condomina/Cuenca 2013, 41). The possibility of neutralisation increases when the demonstrative is in subject position or can be pronominalised by a weak pronoun in Catalan. Overmarking also indicates the existence of different rules than for the use of demonstratives and other referential devices in Catalan and English.

In conclusion, both pragmatics and syntax must be taken into account to describe how demonstratives and other reference devices are used in actual discourse.

4 Subjectivity and modalisation: Modal particles in Catalan

Text is the product of a subject, the addressor, who wants to communicate with another person, the addressee. Sometimes the addressor introduces his or her subjectivity into the text by using modal markers such as modal expressions, predicates and auxiliaries or by resorting to the subjunctive mood, as in (21), where a simple idea such as ‘no acabar la feina’ (‘not to finish work’) is modalised in several ways:⁵

- (21) *La veritat, és possible que no pugui acabar la feina*
 ‘Honestly, it is possible that I may not be able to finish the work’

Among modal markers, modal particles are special in that they lack propositional content. Modal particles convey interactional pragmatic meanings that are often

5 For a general overview of modalisation in Catalan, see Cuenca (2008a, chap. II.2), and volume 42 of the journal *Articles de didàctica de la llengua i de la literatura* (Vilà/Grau 2007).

difficult to describe since they are contextually bound. Catalan has a number of modal particles.⁶ Some of them are not found in other languages, others have counterparts in other languages, although the way they function might differ, whereas others are shared with Spanish, probably because they are calques.

Torrent (2011) offers a systematic account of the most significant Catalan modal particles: *pas*, derived from the homophonic noun *pas* ('step'), *pla*, derived from the adjective *pla* ('flat', 'level'), *rai*, whose origin remains unknown, the particle *si*, an unaccented variant of the affirmative adverb *sí* 'yes', and the particle *que*, corresponding to the general complementiser *que* 'that', when it introduces a question with a falling intonation pattern.⁷

- (22) a. No has *pas* vist una gateta per aquí?
 'You haven't seen a little kitten around here, have you?'
 b. Ara *pla* que hi aniré!
 'Now I will go!'
 c. Vosaltres *rai*!
 'You don't need to worry about a thing!'
 d. Que tens gana! *Si* acabes de dinar!
 'You're hungry? But you've just had lunch!'
 e. *Que* tens hora?
 'Got the time, have you?' [Torrent's (2011) examples and translations]

One of the most idiosyncratic modal particles in Catalan, *pas*, is homophonous with the noun *pas* 'step', which derives from Latin PASSUM. The particle *pas* was allegedly introduced as a means of reinforcing negation in sequences such as *no camin pas* (lit. = 'not walk step'; 'I don't walk even a step') (Espinal 1993, 353). Nowadays it is an adverb of reinforcement used in combination with *no* and it usually follows a verb (23):

- (23) a. No m'ha donat *pas* el llibre
 'He did not *pas* give me the book'
 b. No m'ha *pas* donat el llibre
 'He did not give *pas* me the book'

⁶ Another interesting group of modal markers from a pragmatic perspective are interjections. For a description of Catalan interjections, see Cuenca (2011).

⁷ Torrent (2011) also describes the modal uses of the adverbs modal *ja* 'already', *prou* 'enough', *potser* 'perhaps' or *també* 'too', and mentions *per això* (lit. 'because of that') and certain uses of *ara*, *però*, *cada*, *bé*, and *no*, that could be considered modal but are "difficult to classify [...] as true modal particles" (Torrent 2011, 3.9). See Espinal (2011), in the same volume, for an analysis of the syntactic and semantic aspects of pragmatic particles, including modal ones.

The particle *pas* can also be combined with *no* in configurations related to focalisation (24a) or contrastive constructions (24b) preceding any constituent except for a finite verb form:

- (24) a. Ho ha fet ell, (i) no *pas* jo
 ‘He did it, (and) not *pas* me’
 b. No ho he fet *pas* jo sinó ell
 ‘I did not *pas* do it, but him’

This particle also exists in French and Occitan, but in Central Catalan it exhibits a specific modal use tied to inferencing. In French it is a necessary component of general negation (*Je ne le vois pas*) and has even become the only negative polarity trigger in spoken language (*Je le vois pas*).⁸ Northern Catalan dialects, which are more closely related to French, also share this use of *pas* as a negative particle without any addition of pragmatic value, and the possibility of using the particle without the negative adverb *no* can also be observed in informal language: *Mengi pas; Que vindràs pas a Barcelona?* (GIEC 2016, 1310).⁹

However, in some varieties of Catalan (mainly Central Catalan), *pas* is a modal marker. It behaves as a syntactically optional reinforcement of *no* by which the speaker activates inferential values related to presuppositions (Espinal 1993; 2002; GIEC 2016, 35.4.2.2; Torrent 2011).

“Thus, the speaker’s selection of *no...pas*, instead of *no*, is to be understood as a grammatical sign that the relevance of any given utterance is going to require access to a specific set of propositions which forms part of the most accessible cognitive environment, and *no...pas* is going to either cancel or confirm it. By choosing the complex *no...pas* negative item, rather than the simple *no*, the communicator presents a choice to the hearer, and the hearer has to select the right context for interpretation” (Espinal 1993, 359–360).

The modal particle *pas* can be defined as a reinforcer that activates the polyphonic nature of negation in that it echoes “an opinion attributed to somebody, mainly some

⁸ Interestingly, Hansen (2013) considers the origin of *ne...pas* in French to coincide with the inferential values of *no...pas* in Catalan. The particle *pas* used to be pragmatically marked and syntactically optional in French, as it is now in Catalan. However, it has since become pragmatically neutral and syntactically obligatory, and still is in formal French. The last stage of this evolution is represented by those uses in contemporary French by which *ne* becomes optional and pragmatically marked, whereas *pas* is obligatory and neutral.

⁹ As Torrent (2011, 89) points out, the uses of *pas* exhibit a progression from North to South.: “[...] what starts as a merely semantic-syntactic use (in French) later goes on to play an optional reinforcing role (Catalan in Girona and related dialects) and from there finally has a clearly inferential use (in the Catalan spoken in Barcelona or Vic and related dialects).” In other Catalan-speaking areas, such as the Valencian Country and the Balearic Islands *pas* is not used in spontaneous speech.

thought previously held either by the speaker of the *no...pas* utterance, or by the audience” (Espinal 1993, 359).

Constructions including *pas*, in contrast with those including the general marker *no*, are marked constructions. They activate a different way of processing the utterance, because *pas* triggers an instruction for the hearer to find “contextual assumptions and decide whether the utterance yields to the cancellation of a positive proposition or to the reinforcement of a negative proposition” (Espinal 1993, 359). So both *No m’ha donat el llibre* and *No m’ha donat pas el llibre* refer to the same state of affairs (‘He has not given me the book’), but the version with *pas* implies that the addressor presents the statement as either flouting his or her expectations (‘I expected him to give me the book but he didn’t’) or confirming them (‘As I expected, he did not give me the book’).

The inferential uses of *pas* interact with the type of speech act.

- a) In negative imperative utterances, *pas* reinforces a prohibition. It adds an emphatic or threatening value to the perlocutionary force of the utterance, roughly equivalent to the meaning of the English modal verb *dare*.

(25) No ho facis *pas*
 ‘Do not do it *pas*’

The difference with a simple imperative (*No ho facis*) is that *pas* triggers the inference that negative consequences will arise if the command is not fulfilled. The presupposed content here is that the speaker “takes it for granted that somebody else has an intention” (Torrent 2011, 90) and tries to prevent him or her from taking action.

- b) In questions, *pas* indicates that the speaker expects an affirmative answer, so interrogatives including *pas* can be associated with conformation-seeking questions (cf. Prieto/Rigau 2011).

(26) a. No voldràs *pas* una miqueta de peix?
 ‘Don’t you want *pas* some fish, (right)?’
 b. No deus haver perdut *pas* el mòbil?
 ‘You haven’t lost *pas* your mobile phone, huh?’

In negative interrogatives including *pas* the addressor anticipates a reaction from the addressee and invites him or her to infer that (s)he hopes that the proposition is true (‘I hope that you want some fish’), or fears that it is not true (‘I fear that you have lost your phone’) (Torrent 2011, 89–90). This marked interpretation contrasts with other types of interrogations, as Espinal (1993) shows:

(27) a. Que tens sucre?
 ‘Do you have any sugar?’

- b. No tens sucre?
'Don't you have any sugar?'
- c. No tens *pas* sucre?
'Do you happen to have any sugar?' (examples from Espinal 1993, 365)

In (27a), a question where the particle *que* is optional, the speaker asks the addressee to confirm whether he or she has sugar but has no presuppositions about this fact. In (27b), the speaker has some evidence of the falseness of the assumption and requests confirmation that the addressee does not have any sugar. In (27c), the speaker expects the addressee to have some sugar. Since (s)he has some doubt over this presupposition, the speaker asks whether the proposition is true or false and is ready to make some inferences depending on the answer, "for example, if you don't have any sugar, I'll have to rush to the shopping centre; if you have some sugar, you will probably give me some of it; etc." (Espinal 1993, 365).

- c) In declaratives, *pas* intensifies a previous negation, whether explicit or implicit. Specifically, the particle *pas* cancels "a proposition that is either part of the most accessible context or is an inference deducible from the utterance's context" (Espinal 1993, 354), as shown in (28) and (29), respectively.

- (28) a. En Joan ja no vindrà, a aquestes hores
'John will not come anymore, it is late'
- b. Efectivament, en Joan no vindrà *pas* tan tard
'Sure, John will not come *pas* so late'
- (29) –Em fa l'efecte que en Joan ja deu haver arribat
'I think that John must have arrived'
- No, no ha arribat *pas*. ¿No veus que no hi ha la bossa on ell la sol deixar?
- 'No, he hasn't *pas*. Don't you see that his bag is not where he usually puts it? (examples from Espinal 2002, 2751)

As many of the previous examples show, *pas* is generally embedded in reactive dialogic interventions and indicates that what is said is contrary to someone's expectations.

In conclusion, although Catalan does not have a rich system of modal particles, some of them are used in peculiar ways. This is the case for *pas*, which codifies an instruction to find additional information related to the speaker's expectations and contextual assumptions "either for the reinforcement of a negative proposition, or for a cancellation of a positive proposition, or to certain inferences derived from what the speaker/hearer considers as a desirable thought, or it leads to a prohibition reinforcement, or threat" (Espinal 1993, 367). The correlation *no...pas* is a marked option for negation that adds to the general cancellation of truth-value of the negation "an overtone of rejection or confirmation of an expectation" (Espinal 1993, 361).

5 Discourse marking: Parenthetical and pragmatic connectives

Connectives and discourse markers in general help to tie together different parts of the sentence and portions of the text (see Cuenca 2008a, chap. IV). At the textual level, they contribute to discourse organisation.

Catalan has a complex system of connectives, including conjunctions and parenthetical connectives, among other less fixed or prototypical markers (see Cuenca 2002; 2006; 2008a; Castellà 2004; González 2004; GIEC 2016, chap. 25). At the text level, only a few conjunctions can act as connectives, among which the most frequent are *i* ‘and’ and *però* ‘but’ (30):

- (30) València és una ciutat plena de monuments històrics i llocs per visitar.
Però cal no oblidar que és també una destinació de platja
‘Valencia is a city full of historical monuments and places to visit. But it is important not to forget that it is also a beach destination’

Independent sentences or groups of sentences are frequently linked by parenthetical connectives. Parenthetical connectives are appositional, syntactically detached words or fixed phrases roughly equivalent to conjunctions from a functional point of view.¹⁰

- (31) Volia venir al viatge. *Tanmateix*, s’ha hagut de quedar
‘She wanted to come to the trip. However, she had to stay’

Their general meanings can be grouped in four types, which include several more specific values, as indicated in Table 2, which includes the most frequently used parenthetical markers in Catalan.

Table 2: Main Catalan parenthetical markers.

General meaning	Specific meaning	Markers
a. Addition	Continuity ‘and’	a continuació; a més; d’altra part; doncs (bé), llavors/aleshores
	Intensification ‘moreover’	a més (a més), més encara
	Sequencing ‘on the one hand/on the other hand’	d’una banda...de l’altra per una banda...per altra banda;

¹⁰ For a definition of these categories, see Cuenca (2013).

General meaning	Specific meaning	Markers
	'firstly, secondly...'	per començar, d'entrada, d'antuvi, en primer lloc, en segon lloc, etc.;
	'finally'	finalment, per fi, per acabar
	Digression 'by the way'	per cert
	Generalisation 'generally'	en general, generalment
	Specification 'specifically', 'indeed'	en concret; fet i fet, de fet
	Amplification 'in fact'	de fet, en efecte, efectivament, certament, per descomptat
	Equiparation 'similarly'	així mateix, semblantment, igualment
b. Disjunction	Reformulation 'that is' 'or rather'	és a dir; això és; o sigui/siga/sia; més ben dit; ras i curt; més aviat.
	Exemplification 'for example'	per exemple, a tall d'exemple, així (per exemple)
	Summary 'in sum'	en resum, en síntesi, en suma, comptat i debatut
c. Contrast	Opposition 'however' 'otherwise'	tanmateix; en canvi; ara (bé); bé; altrament; si no; en cas contrari
	Concession 'nevertheless'	no obstant això/això no obstant; amb tot, (amb) tot i amb això; tot i així/això; malgrat tot; de tota manera; en tot cas; en qualsevol cas; sigui com sigui/siga com siga; al capdavant
	Restriction 'rather'	si més no, almenys, més aviat
	Refutation 'on the contrary'	(ans) al contrari, ben al contrari, per contra
	Contraposition 'as a matter of fact'	en realitat; de fet, fet i fet; ben mirat
d. Consequence	Consequence 'so, 'as a consequence'	(així) doncs, llavors/aleshores, per tant, així, per consegüent
	Conclusion 'in conclusion' 'all in all'	en fi, en conclusió, en definitiva; per concloure; al capdavant; fet i fet; comptat i debatut

Parenthetical connectives can indicate connection both at the text level (inter-sentential use) and at the sentence level (intra-sentential use) either on their own or following a conjunction, as the following examples show:

- (32) Ha fet tard i, *a més*, no portava els documents
'He was late and, in addition, he was not bringing the documents'
- (33) Ha arribat a les onze, *és a dir*, ha fet tard
'She arrived at eleven, that is, she was late'
- (34) El professor d'anglès és molt actiu. La professora d'història, *en canvi*,
només llegeix el llibre de text
'The English teacher is very active. The History teacher, on the contrary,
only reads the textbook'
- (35) No hi ha proves contra ell. *Per tant*, l'han deixat lliure
'There is no evidence against him. As a consequence, he's been released'

The previous examples illustrate the four general meanings that can be expressed at the text level: *a més* indicates addition, *és a dir* is a reformulation that introduces an alternative formulation of the intended meaning, *en canvi* opposes two independent sentences, and *per tant* introduces a consequence.

Along with parenthetical connectives, there is a group of appositional markers typically used in oral texts and in dialogue such as *bé* 'well', *miri/mira* 'look', *a veure/aviam* 'let's see', *escolti/escolta* 'listen', *home* (literally 'man') and *dona* (literally 'woman').

- (36) Em va passar una cosa molt estranya. *Mira*, anava de passeig i se'm va
acostar un home tot vestit de negre...
'Something very strange happened to me. Look, I was taking a walk and a
man all dressed in black came to me...'
- (37) –Li donaràs la clau al teu germà, no?
'You will give your brother the key, won't you?'
–*Bé*, no sé si podré.
'Well, I don't know if I can'

In (36) *mira* prefaces the beginning of a narrative after an introductory sentence and tries to capture the hearer's attention. In (37) *bé* prefaces a response that is contrary to the previous interlocutor's presupposition and thus announces some kind of disagreement.

These appositional and syntactically detached markers can be grouped together as *pragmatic connectives*.¹¹ Pragmatic connectives typically preface interventions, turns or units within turns while indicating structural and interactional meanings.

Since an exhaustive account of the functions of all pragmatic connectives would exceed the limits of this chapter, I will focus on the marker *home*, according to the analysis included in Cuenca/Torres (2008) and summarised in Cuenca (2013).

The noun *home* ('man') – and less frequently its feminine form *dona* ('woman') – is often used as a marker in Catalan conversation. As a pragmatic connective, *home* introduces turns and utterances inside turns, as illustrated in (38, from Cuenca/Torres 2008):¹²

(38)

EUU:	no saps ni obrir [una cerve]sa\	you can't even open a beer
PUY:	[ah sɪ̃] (soroll de l'obridor)	oh, yes (sound of an opener)
MJJ:	<i>home</i> \ mira\ obrir_ obrir_ el que és obrir_ <i>home</i> 'man', look, opening_ opening_ opening a una cervesa_ no se'm dóna bé\ (COC05, beer is not something I'm very good at 159–170) [M-W]	

In (38) a man (MJJ) responds to a woman's (EUU) reproach. The use of *home* to introduce his turn softens his admission that he is not able to open a beer can.

As a marker, *home* (and *dona*) is almost completely fixed and admits no specifiers or complements. It looks like a vocative but has completely lost its referential meaning and part of its appellative force. The marker *home* shows a strong tendency to invariability. It has no plural and the masculine form, *home*, can be used to refer to a plural addressee. Moreover, *home* is frequently used to address both women and men, whereas *dona* can only be addressed to a woman and it is mainly found in final position, where its source meaning is more persistent.¹³

The marker *home* is used in conversation as a solidarity marker that identifies the interlocutors as people that know each other and are engaged in a largely informal

11 On Catalan pragmatic markers, including pragmatic connectives, see Cuenca (2006; 2007); González (2004); and also the works by Cuenca/Marín (2001); Montolío/Unamuno (2001); Marín (2005a; 2005b) on markers derived from perception verbs; Cuenca/Torres (2008) on *home/dona*; and Cuenca (2008b) on the Catalan counterparts of the English discourse marker *well*.

12 The examples come from *Corpus de Conversa Col·loquial* (COC) (Payrató/Alturo 2002). Since the examples correspond to informal conversation, the glosses will be quite literal. In the following examples, a final coda identifies the interlocutors' gender (woman, W, man, M).

13 In the COC corpus, *home* ('man') is addressed to a woman in many of the examples (78 out of 133 cases), 41 of which are used by a man addressing a woman and 37 are used by a woman addressing another woman. There is only one case of *dona* in the corpus, although 35 of the participants were women.

conversation. By using *home* (or *dona*), the addressor conveys modal meanings related either to the mitigation of a disagreement (or partial agreement) or to emphasis.

In initial or left-periphery position *home* is often used to soften an utterance that indicates disagreement or partial agreement (39):

(39)

PEI:	lo que passa es que vosaltres sí que podeu posar-hi una_ un: anunci (. 0.12) o no\	the point is that you can place an ad, can't you?
RRR:	(.0.26) <i>home</i> \ però si ja ho tenim dit a tothom\ (COC08, 34–39), [W-W]	<i>home</i> 'man', we've already told everyone

In (39) *home* introduces an utterance that disagrees with PEI's proposal ('you can place an ad') and argues against it ('since we've already told everyone there is no point in placing an ad').

In final (or intermediate) position, *home*, and the less frequently used variant *dona*, are used to reinforce an utterance, often a command (40), or a polarity item, either positive or negative (41):

(40)

VIE:	divendres_ la truquem\	we'll call her on Friday, huh?
ALL:	(.. 0.34) eh:/	uh?
JON:	al vespre truqueu\ i ja està [<i>home</i> \]	you call her in the evening and that's it, <i>home</i> 'man'
ALL:	[<i>va</i> \] (COC10, 1252–1257) [M-W]	OK

(41)

ANA:	{(??) voleu aigua\}	do you want some water?
COC:	no: <i>dona</i> \	no, <i>dona</i> 'woman'
COC:	espera que acabin de [xxx\ (COC04, 210–212) [W-W]	wait until they finish...

In summary, *home* is a multifunctional marker. Its meaning and behaviour are conditioned by its position. In initial position, the connective nature is more prominent and the marker mitigates a (total or partial) disagreement. In final position, *home* (or *dona*) is generally an emphasis marker but it can also mitigate a command.

6 Conclusion

Pragmatics and text linguistics are vast fields of analysis covering crucial concepts that help us to understand how communication takes place. In this chapter, I have presented several phenomena that exhibit particular behaviour in Catalan, providing examples from several case studies that highlight the interaction between pragmatics, grammar and text. First, the use of the definite article with personal names illustrates the importance of shared knowledge and presuppositions in discourse construction. Second, Catalan demonstratives show great variation in their use and interaction with other reference devices, an observation which becomes even clearer from a cross-linguistic perspective. Third, modal particles in general, and the modal particle *pas* in particular, stand out as a means of expressing subjective meanings and modalisation. Finally, this chapter provided a brief description of discourse markers, and specifically parenthetical markers and pragmatic markers, and their use as connective devices that help create text coherence bracketing units of talk while expressing pragmatic meanings.

7 Bibliography

- Alturo, Núria (2003), *La pragmàtica en la tradició sociolingüística catalana*, Noves SL, Hivern 2003, <http://cultura.gencat.es/llengcat/noves/edit-hiv03.htm> (last accessed: 15.032017).
- Alturo, Núria (2011), *Pragmàtica i anàlisi del discurs*, Treballs de Sociolingüística Catalana 21, 29–41.
- Badia i Margarit, Antoni Maria (1994), *Gramàtica de la llengua catalana. Descriptiva, normativa, diatòpica, diastràtica*, Barcelona, Enciclopèdia Catalana.
- Bassols, Margarida (2001), *Les claus de la pragmàtica*, Vic, Eumo.
- Bassols, Margarida (2003), *Pragmàtica i anàlisi del discurs*, in: Noves SL, Hivern 2003, <http://cultura.gencat.es/llengcat/noves/edit-hiv03.htm> (last accessed: 15.032017).
- Beaugrande, Robert de/Dressler, Wolfgang (1981), *Introduction to Text Linguistics*, London, Longman.
- Boix, Emili/Payrató, Lluís (1997, 1995), *An Overview of Catalan Sociolinguistics and Pragmatics*, Catalan Review 9/2, 317–403.
- Brucart, Josep Maria (2002), *Els determinants*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, Barcelona, Empúries, 1435–1516.
- Castellà, Josep Maria (1992), *De la frase al text. Teories de l'ús lingüístic*, Barcelona, Empúries.
- Castellà, Josep Maria (2004), *Oralitat i escriptura. Dues cares de la complexitat del llenguatge*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Cuenca, Maria Josep (2002), *Els connectors textuais i les interjeccions*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, Barcelona, Empúries, 3173–3237.
- Cuenca, Maria Josep (2006), *La connexió i els connectors. Perspectiva oracional i textual*, Vic, Eumo.
- Cuenca, Maria Josep (2007), *La pragmàtica en la gramàtica*, Barcelona, Institut d'Estudis Catalans.
- Cuenca, Maria Josep (2008a), *Gramàtica del text*, Alzira, Bromera.
- Cuenca, Maria Josep (2008b), *Pragmatic Markers in Contrast. The Case of "well"*, Journal of Pragmatics 40, 1373–1391.
- Cuenca, Maria Josep (2011), *Catalan Interjections*, in: Lluís Payrató/Josep Maria Cots (edd.), *The Pragmatics of Catalan Language*, Berlin/Boston, De Gruyter Mouton, 173–211.

- Cuenca, Maria Josep (2013), *The Fuzzy Boundaries Between Discourse Marking and Modal Marking*, in: Liesbeth Degand/Bert Cornillie/Paola Pietrandrea (edd.), *Discourse Markers and Modal Particles: Categorization and Description*, Amsterdam/Philadelphia, Benjamins, 191–216.
- Cuenca, Maria Josep/Marín, Maria Josep (2001), *Verbs de percepció gramaticalitzats com a connectors. Anàlisi contrastiva espanyol-català*, in: Ricardo Maldonado (ed.), *Estudios cognoscitivos del español*, Revista Española de Lingüística Aplicada (monographic issue), 215–223.
- Cuenca, Maria Josep/Torres, Marta (2008), *Usos de “hombre/home” y “mujer/dona” como marcadores del discurso*, *Verba* 35, 235–256.
- Espinal, M. Teresa (1993), *The Interpretation of “no...pas” in Catalan*, *Journal of Pragmatics* 19, 353–369.
- Espinal, M. Teresa (2002), *La negació*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 3, Barcelona, Empúries, 2727–2797.
- Espinal, M. Teresa (2011), *Pragmatic Particles at the Syntax-Cognition Interface*, in: Lluís Payrató/Josep Maria Cots (edd.), *The Pragmatics of Catalan Language*, Berlin/Boston, De Gruyter Mouton, 49–79.
- GIEC = IEC (2016), *Gramàtica de la llengua catalana*, Barcelona, IEC.
- González, Montserrat (2004), *Pragmatic Markers in Oral Narrative. The Case of English and Catalan*, Amsterdam/Philadelphia, Benjamins.
- Halliday, M. A. K./Hasan, Ruqaiya (1976), *Cohesion in English*, London, Longman.
- Hansen, Maj-Britt Mosegaard (2013), *The History of Negation in French*, in: David Willis/Christopher Lucas/Anne Breitbarth (edd.), *The History of Negation in the Languages of Europe and the Mediterranean*, vol. 1: *Case Studies*, Oxford, Oxford University Press, 51–76.
- Marín, Maria Josep (2005a), *Marcadors discursius procedents de verbs de percepció: argumentació implícita en el debat electoral*, *Quaderns de Filologia*, Annex 59, València, Universitat de València.
- Marín, Maria Josep (2005b), *Gramaticalització i funció discursiva dels verbs de percepció*, *Caplletra* 38, 47–71.
- Montolí, Estrella/Unamuno, Virginia (2001), *The Discourse Marker “a ver” (Catalan, “a veure”) in Teacher-Student Interaction*, *Journal of Pragmatics* 33, 193–208.
- Nogué, Neus (2008), *La dixi de persona en català*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Nogué, Neus (2011), *Person Deixis in Catalan*, in: Lluís Payrató/Josep Maria Cots (edd.), *The Pragmatics of Catalan Language*, Berlin/Boston, De Gruyter Mouton, 115–144.
- Nogué, Neus (2015), *Catalan*, in: Konstanze Jungbluth/Federica Da Milano (edd.), *Manual of Deixis in Romance Languages*, Berlin/Boston, De Gruyter, 206–239.
- Payrató, Lluís (2002), *L'enunciació i la modalitat oracional*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, Barcelona, Empúries, 1149–1220.
- Payrató, Lluís (2016), *Estudis de pragmàtica i anàlisi del discurs sobre la llengua catalana, un repàs de l'etapa 1997–2012*, *Estudis Romànics* 38, 55–88.
- Payrató, Lluís/Alturo, Núria (edd.) (2002), *Corpus oral de conversa col·loquial. Materials de treball*, Barcelona, Publicacions de la Universitat de Barcelona.
- Payrató, Lluís/Cots, Josep Maria (edd.) (2011), *The Pragmatics of Catalan Language*, Berlin/Boston, De Gruyter Mouton.
- Prieto, Pilar/Rigau, Gemma (2011), *Prosody and Pragmatics*, in: Lluís Payrató/Josep Maria Cots (edd.), *The Pragmatics of Catalan Language*, Berlin/Boston, De Gruyter Mouton, 17–47.
- Ribera i Condomina, Josep (2012), *La cohesió lèxica en seqüències narratives*, Alacant/Barcelona, Institut Interuniversitari de Filologia Valenciana/Publicacions de l'Abadia de Montserrat.
- Ribera i Condomina, Josep/Cuenca, Maria Josep (2013), *Use and Translation of Demonstratives in Fiction. A Contrastive Approach (English-Catalan)*, *Catalan Review* 27, 27–49.
- Rigau, Gemma (1981), *Gramàtica del discurs*, Bellaterra, Universitat Autònoma de Barcelona.

- Torrent, Aina (2011), *Modal Particles in Catalan*, in: Lluís Payrató/Josep Maria Cots (edd.), *The Pragmatics of Catalan Language*, Berlin/Boston, De Gruyter Mouton, 81–113.
- Vilà, Montserrat/Grau, Maria (edd.) (2007), *Modalització i cortesia lingüística*, *Articles de Didàctica de la Llengua i la Literatura* 42, 5–7 (special issue).
- Wheeler, Max/Yates, Alan/Dols, Nicolau (1999), *Catalan: A Comprehensive Grammar*, London/New York, Routledge.

7 Lexicon

Josep Martines

7.1 General Lexicon

Abstract: This chapter describes the basic historical components of the Catalan lexicon. Here, the author analyzes processes of change and variation that can be observed in the Catalan lexicon, paying special attention to those that should be reviewed and subjected to critique. The main part of this chapter discusses lexical-semantic variation and change. Specifically, from a diachronic perspective, the author describes the mechanisms of semantic change, the factors that may reduce a certain word's usage, as well as mechanisms and factors of word creation. From a territorial or dialectal perspective, the author offers a characterization of the Catalan lexicon from a contemporary and historical lexical-semantic perspective. Next, the author provides historical and contemporary materials to aid in the description of social and functional variation in the Catalan lexicon. To conclude, this chapter addresses current and future challenges faced by the Catalan lexicon.

Keywords: general lexicon, language strata, lexical-semantic diachrony, dialectal, functional and social change and variation

1 The Catalan lexicon: overview

Following studies published by Colón (1976) there is widespread consensus in considering Catalan, with regard to its lexicon, to be part of the Gallo-Romance family, together with Occitan and French. It is certainly difficult to ignore the extent to which their basic vocabulary coincides: *matí* 'morning', *demà* 'tomorrow', *finestra* 'window', *menjar* 'to eat' clearly correspond with *matin*, *deman*, *fenèstra*, *manjar*, in Occitan; and with *matin*, *demain*, *fenêtre*, *manger*, in French. These correspondences were even greater in the Middle Ages, before Catalan underwent a so-called *hispanization* process, in the wake of its reorientation from Occitania towards the Iberian Peninsula and the Mediterranean: *frare* 'brother', *sor* 'sister' or *jorn* 'day' (close to *fraire*, *sòr* or *sòrre*, *jorn* or *frère*, *sœur*, *jour*) gave way to *germà*, *germana* and *dia*.

Throughout its history, the Catalan lexicon has experienced significant changes, associated with both internal and external conditions. In the first case, there is still much to be done, most notably the application of contemporary corpus analysis methodologies, before the field has been sufficiently explored. In the second case, diachronic sociolinguistics has provided a way to gauge to what extent the socio-

<https://doi.org/10.1515/9783110450408-012>

political context and its contingencies have marked the present state and the evolution of the Catalan lexicon. From a flourishing language used in the administration, science and literature, used by both the common people and the upper classes as well as in ecclesiastical sermons, commerce and economic activity in the Middle Ages, Catalan underwent a progressive minoritization from the Modern Age onwards to the present, with the corresponding effects of this process on the internal development of the language.

The intense cultivation of Catalan over history has put within our reach an immense wealth of literary and non-literary texts, whose study may serve to trace this language's travels through time and space. Concerning the historical study of the lexicon, recently developed digitalized corpora (CTILC, CICA, CIMTAC) are crucial aids in this project, as well as large historical and etymological dictionaries (DCVB, DECat, FEW, *D'Aguiló*, DCECH), lexicographic compilations (TLV, LEXDIALGRAM), and numerous author and document monographs. To analyze the current state of the language, we may use language atlases (ALDC, ALDT, PALDC), the countless dialect monographs that have appeared recently, and even oral corpora (CCCUB).

Within the scope of this chapter, we will attempt to offer a brief description of the components of the Catalan lexicon (section 2); the processes of variation and change that can be observed in the Catalan lexicon (3), with a focus on the temporal (3.1), the dialectal (3.2), the social and the functional dimensions (3.3); and we shall conclude with a characterization of the challenges facing Catalan today and in the future (4). This way, we aim to combine the analysis of internal and external factors that may explain the evolution of the Catalan lexicon.

2 Historical components in the Catalan lexicon

In any language, the study of the origins of its lexicon leads invariably to the history of its speakers: the communities that have coexisted with it, the cultural influences it has been exposed to, the sociolinguistic contexts that have emerged in each of its historical periods. Despite the rigidity of the *substrata* metaphor (coined in the 19th century by Graziadio Isaia Ascoli) and its lack of proper sociolinguistic analysis, it may serve to conceptually represent the historical process of sedimentation, and the influences between languages and cultures that have accumulated in "layers" over a given territory.

The study of the origins of the Catalan lexicon has a solid and significant tradition. Briefly, its foundations are a) the portentous historical and etymological dictionaries (see above), b) the accounts of historical grammar and language history (Moll 1952; Badia i Margarit 1981; Sanchis 1980; Nadal/Prats 1982–1996; Duarte/Alsina 1984; Bruguera 1985; Ferrando/Nicolás 2011; Batlle et al. 2016), c) the prolific work done by Coromines, Veny, Badia, Colón, Ferrando, Miralles, Moran, Casanova, Colomina, etc., both in text editions and text analysis, as well as the individual biographies of many

words, and d) currently, digitalized text corpora (see above) and the corresponding monographic studies on lexical history and diachronic semantics made possible by these corpora (cf. 3.2). In what follows we will briefly trace the lexical strata of Catalan, with special attention to aspects recently subjected to criticism and revision; we will omit personal names and place names (▷15 Onomastics: Personal Names and Place Names), at times the only trace that remains of some ancient languages.

Some Catalan words may have originated in periods preceding the arrival of Latin to the Iberian Peninsula (more precisely, the northeastern region), dated in 218 BC in the context of the Second Punic War between Carthage and Rome. From this era, much remains still in the shadows, including an exact account of the languages spoken in the territory, their typological relationships, sociolinguistic situation, and any lexical items that may have survived into Hispanic Latin. For instance,

- a) Within the Indo-European context, which includes language strata previous to Indo-European immigration to the Iberian Peninsula, (ca. the 1st millennium BC), there is still some uncertainty regarding the language conditions in the Pre-Roman western Mediterranean and, more concretely, of the Iberian Peninsula, south of Gaul and Northern Africa, as well as uncertainty regarding which relationships exist between attested languages (apart from Old Basque, more properly named Aquitanian, and Iberian). More concretely, the language of the Iberians is not well known (in fact, most of their numerous inscriptions have been deciphered, but cannot be interpreted), neither are the relationships between Old Basque or Aquitanian (on both sides of the central and western Pyrenees) and the language of the Iberians (whose culture seems to have flourished in the eastern peninsula between the 6th and 3rd centuries BC).
- b) Within the Indo-European context, the identity (Celtic or otherwise) of various migratory movements toward the Iberian Peninsula has not yet been ascertained, neither have their linguistic attestations been fully understood. In this context, Joan Coromines' theory of a *sorothaptic* language (associated with the *Urnenfelder* culture) attempts to explain the origins of Catalan lexical elements of a non-Celtic, Indo-European origin, although it has not achieved broad consensus. Let us observe that, in both cases, a) and b), it is often difficult to decide whether a given word originated directly in the substratum, or whether it arrived by way of a third language.
- c) There are still very many words attributed to pre-Roman times without specific evidence to link them to a language.¹ The following is a brief sample of allegedly pre-Roman, Catalan words, based on the work of Coromines. i. Non-Indo-European words, from Basque or Iberian: *caparra* 'leech', *estalviar* 'save', *socarrar* 'burn', *gavarrera* 'wild rose'; of unclear pre-Roman origin: *bassa* 'raft', *barraca*

¹ For more on this issue, see Adiego (2002), Correa (2005), Gorrochategui (2002), Sánchez Moreno (2007), Villar (2002; 2008).

‘shack’, *carabassa* ‘pumpkin’, *carrasca* ‘holm oak’. ii. Indo-European words identified as non-Celtic: *cabàs* ‘wicker basket’, *carant* ‘water current on a rocky surface’, *cabanya*, *-ana* ‘cabin, hut’, *tancar* ‘to close’. iii. Words identified as Celtic: *bresca* ‘honeycomb’, *dalla* ‘scythe’, *galta* ‘cheek’, *bruc* ‘heather’, *escombrar* ‘to sweep’, *volva* ‘snowflake’, *banya* ‘horn’, *agafar* ‘to take’, *basca* ‘anguish, nausea’.

Catalan contains an extensive list of words of Germanic origin (Coromines 1952; Bruguera 1985; Jaime 2015; 2017). A subset of these may have entered Catalan by way of spoken Latin, as a consequence of the long interaction (both in war and peace) between Rome and Germanic tribes. This is likely the case with *bandera* ‘flag’, *braó* ‘part of the arm located between elbow and shoulder’, *sabó* ‘soap’, *blau* ‘blue’, *bru* ‘brown’, *roba* ‘clothes’, *guerra* ‘war’. The disintegration of the Roman Empire (5th century) and the expansion of Germanic groups intensified the penetration of Germanic lexical elements into Latin, which was already starting to morph into the various Romance languages. Catalan inherited many words of Gothic and Frankish origin during its brief period as a part of the Visigoth Kingdom (5th–8th centuries) and, later, of the Frankish Empire (8th–9th centuries), following the Marca Hispanica. Of these words, some may well have been present already in Vulgar or Late Latin; in the latter case, the words probably arrived into Catalan during the Carolingian period and the expansion of feudalism in Catalonia during the High Middle Ages, or by way of Occitan. Words of Gothic origin are *conrear* ‘farm, cultivate’, *melsa* ‘spleen’, *treva* ‘truce’, *sàrria* ‘long pannier, basket’, *ranc* ‘limp’, *fang* ‘mud’. Words of Frankish origin are *jaquir* ‘to allow, to permit’ and ‘leave, abandon’ (Martines/Montserrat 2014), *òbila* or *òliba* ‘owl’, *bugada* ‘laundry, washing’, etc.

When Arabic entered the Iberian Peninsula in the 8th century, it triggered an extended period of contact with developing Romance languages that ended in 1609 (after the expulsion of Muslims). In the case of Catalan, though Arabic words are fewer than in other Ibero-Romance languages such as Castilian and Portuguese, there is still a significant corpus of diverse semantic domains (Coromines 1977; Corriente 1999a; 1999b; Barceló 1984; 2011): *albergínia* ‘aubergine’, *aljub* ‘cistern, tank’, *atzucac* ‘cul-de-sac’, *matalaf* or *matalàs* ‘mattress’, *safareig* ‘basin’, ‘tank’, *sèquia* or *síquia* ‘ditch’, *talaia* ‘watchtower’, etc. Southern Catalan displays still more words of Arabic origin; some more examples are found in 3.2.

There has been some speculation about the influences in Catalonia and the Valencian region of Mozarabic, or Andalusí Romance, that is, ‘the variety or varieties of Romance maintained for at least some time in al-Andalus’ (Corriente 2001). This speculation has been due, most particularly in the Valencian region, to political and ideological reasons. Toponymy is a useful source of words of Andalusí Romance origin, as their oral transmission (inter-generational and among different languages) tends to be conservative. In contrast, Catalan lacks any solid lexical attestations that may be interpreted as inherited from Latin dialects spoken in Muslim territory (after undergoing minoritization and language shift processes in contact with Arabic). The

following words that are attributed to Mozarabic should in fact be explained as Aragonese (*bresquilla* or *pres-* ‘peach’), Italian (*orxata* ‘almond milk’; *xitxarel-lo* ‘young lad’), or Castilian (*mossiguello* ‘bat’, *torondo* ‘swelling’) in origin.²

Less attention has been paid to the relationship between Aragonese and Catalan. Together with Latin, these two languages became official in the administration of the Catalan-Aragonese Crown and its neighboring realms in the 12th century, after the County of Barcelona united with the Kingdom of Aragon, and Catalan expansion changed its orientation towards the Iberian Peninsula and the Mediterranean (away from Occitania after their defeat in Muret in 1213). As the Crown advanced into the Muslim south, Western Catalan expanded its contact with Aragonese. The Aragonese played a vital role in populating the new kingdom of Valencia (13th century) after it was conquered away from Muslim powers. Catalan and Aragonese populations have maintained reciprocal relations in commerce, livestock and immigration throughout the Middle Ages. These relationships explain some coincidences in linguistic traits between Catalan (especially the western variety, including the Valencian region) and Aragonese, intensifying toward eastern Aragon. Likewise, these relationships explain the presence of words of Aragonese origin (or words that define an area of lexical continuity with Aragonese) in the Catalan domain, intensified in the west and Valencia. Often, these words expand into Castilian varieties on the east of the peninsula, whereby they display varying territorial distributions. For instance: *bresquilla*, *pres-* (see above), *catxirulo* ‘kite’, *gemecar* ‘moan’, *corder* ‘lamb’, *mortitxol* ‘young toddler’, *lligallo* ‘a type of road’, *fardatxo* ‘lizard’, *meló d’aigua* ‘watermelon’, *brosquil* ‘dense forest’, etc. See Veny (2002; 2011), Martines (2002; 2009; 2012b).

Colón (1976) showed the extent to which Spanish contact influenced Catalan during the Middle Ages. As we stated earlier, Catalan drew closer to the Hispanic world (*hispanització*) during the latter centuries of the Middle Ages, while it still maintained an essentially Gallo-Romance character. This tendency expressed itself in the promotion of words formally closer to Spanish, and the inclusion of loan words from Spanish. The political changes that took place during the Modern Age had sociolinguistic consequences that intensified the presence and status of Castilian in the Catalan domain: starting in the late 16th–17th century (the *Siglo de Oro*), the spread and prestige of Castilian grew and the Hispanic monarchy was strengthened; it imposed an absolutist, centralist model of governing, as embodied in the decrees of Nueva Planta, following the War of the Spanish Succession (18th century).

Digital text corpora and close reading of texts of the time provide ample evidence for the intensification of the penetration of the Castilian lexicon in every register, both formal and informal, as well as in every region, even Northern Catalonia (not incorporated into France until 1659 with the Treaty of the Pyrenees, or in l’Alguer

² More on Andalusí Romance and Catalan can be found in Colón (1997), Colón/Ferrando (2011), Barceló (2011), Veny (2011), Martines (2011; 2012b).

(Sardinia), also subjected to Spanish administration. Later, French and Italian were decisive influences in the latter two). Factors such as the spread of Castilian as the language of the Crown of Castile, the prestige of Castilian literature, the popularization and activities of traveling theater companies (Fuster 1976) and church preaching in Castilian (especially in the Valencian region; Pitarch 2001) promoted the familiarity of the speakers of Catalan with Castilian. Add to this the forced imposition of Castilian after the War of the Spanish Succession and the actions of the government to oppose the public use of Catalan, as well as its use in elementary and higher education during the 19th century. The bilingualization of all Catalan speakers was completed during the 20th century, and it was achieved by means of the imposition of Castilian in the administration, schools, developing mass media, and most other public domains. This context had a negative effect (though in some regions more than others) on the sociological status of Catalan and its ability to linguistically integrate the numerous immigrants, who arrived in powerful waves into the Catalan-speaking domains in the second half of the 20th century and afterwards. The explicit and direct actions undertaken in the dictatorship of General Francisco Franco (1936–1975) proved decisive, though even successive democratic regimes have tended to support keeping Castilian in a privileged position. Since then, Castilian has progressively integrated into everyday life, into the language and cultural landscape of the speakers of Catalan, and it has functioned as a means of transmission and categorization of cultural innovations.

The normalization process of Catalan, which solidified at the end of the 19th century and the first decades of the 20th century in the work of Pompeu Fabra and the Institut d'Estudis Catalans (working as an academy of the Catalan language), emerged around the task of de-castilianizing the lexicon and restoring its authentic lexical elements, and to further generate original labels for cultural innovations. This task implies attempting a balance act between keeping the originality of the standard model and its distance to the Castilian language, on the one hand, and avoiding an archaizing model, excessively distanced from real usage (Lamuela/Murgades 1984; Rafel 1996; Veny 2007).

One result of the process described above is the emergence of a lexical repertoire in contemporary, normative Catalan that originated in Castilian. Some of them entered Catalan in the Middle Ages: *amo* 'master' (from Cast. *amo*), for *senyor* or *mestre*; *xop* 'poplar' (from Cast. *chopo*), for *pollancre* (see below); *esmorzar* or *al-* 'meal at mid-morning'; *boda* 'wedding' (from Cast. *boda*), for *noces*. Starting in the Modern Age, Castilian vocabulary entered Catalan more profusely, and some words became part of the normative vocabulary: *preguntar* 'to ask' (from Cast. *preguntar*), for *demanar*; *buscar* 'to search for' (from Cast. *buscar*), for *cercar*; *quedar* 'to remain' (from Cast. *quedar*), for *restar*, *romandre*; *queixar* 'to complain' (from Cast. *quejar*), for *plànyer*, *clamar*; *hisenda* (from Cast. *hacienda*); *estrella* 'star' (from Cast. *estrella*), for *estel*, *estela* or *estrela*; *enfadar* (from Cast. *enfadar*), for *enutjar*. A significant subset of the corresponding Catalan words has survived in a few territories or only in formal

registers, and act as markers of high formality, or else appear in phraseology, sometimes without being understood by speakers who use them. It is interesting to note the presence in Catalan of a Castilian vocabulary that is no longer frequent in that language: *amoïnar* (from Cast. *amohinar* ‘to worry’), *antuvi* (used in the loc. *d’antuvi* ‘in principle’; from Cast. *antuvio* ‘hasty or pre-emptive action’), *aïna* ‘easily’ (also in the loc. *més aïna* ‘rather’; from Cast. *ahina* ‘easily’).

Parallel to this, as we shall see in 4.2, the dissemination process of normalized Catalan has allowed for the recovery into general usage of words that had been replaced by Castilian terms; even the elaboration of new proposals to name new cultural concepts, which also established themselves in general usage, has begun to take place.

Castilian interference has not only affected the purely lexical domain: it has also eroded Catalan semantics. With some territorial differences, we may observe a tendency to adapt the use of verbs *ser* and *estar* or *anar* and *venir* to Castilian usage (*ser*, *estar* ‘to be’; *ir* ‘to go’, *venir* ‘to come’); dismantling the distinctions between *eixugar* ‘to cause a wet object to lose by evaporation whatever water is covering or soaking it’ and *assecar* ‘to cause an object to lose the humidity it contains’, or *capsa* ‘box of small dimensions’ and *caixa* ‘large box’ and favour *assecar* (or *secar*) or *caixa*, closer to Castilian *secar*, and *caja* for both meanings. The use of the Catalan word to name concepts that speakers cognitively associate with past and traditional mental spaces has repercussions in all varieties of Catalan. For example, *atún* ‘tuna’ (from Cast. *atún*) may designate ‘canned tuna’ while *tonyina* ‘tuna’, will refer to fished, unprocessed fish; *lejía* ‘lye’ (from Cast. *lejía*) is bought at the store but *lleixiu* is homemade. There are many more examples, and some have entered the normative dictionary. For instance, *jupetí* ‘doublet’ as ‘*armilla de pagès*’ (‘farmer’s doublet’): the word of Cast. origin *armilla* is presented as a generic term to define the properly authentic term, which has been relegated to a marginal usage, marked as traditional.

Contact between languages often has consequences for both. The Middle Ages witnessed how a few important Catalan words entered the Castilian lexicon (*papel* ‘paper’, from Cat. *paper*; *timonel* ‘helmsperson’, from Cat. *timoner*; *reloj* ‘clock’, from Cat. *rellotge*). Likewise, the eastern varieties of Castilian integrated an important amount of words of Catalan origin. This question has become more relevant in the context of newly elaborated Castilian digital corpora and linguistic atlases.

As stated above, Occitan has a very high affinity with Catalan, indeed most of the general vocabulary in both languages coincides. This affinity was even more intense during the Middle Ages, before Catalan reoriented itself towards the Hispanic world; geographic continuity and human contact explain the presence of a great deal of words of Occitan origin in Northern Catalan. The prestige of troubadour poetry promoted the integration of Occitan literary terms such as *pretz* ‘merit’ or *razó* ‘part of a troubadour poem in which the reasons for its composition are exposed’. Cultural influence, coexistence, commerce and population flows (even during the Modern Age) may explain the transit of the following Occitan words into Catalan: *ambaixada*

'embassy', *bacallà* 'cod', *arenc* 'herring', *imaginaire* 'a sculptor of sacred images' and even morphological-derivational elements such as the suffix *-aire* (Martines 1997). There is still a wide margin of study in this terrain: words such as *punxa* 'thorn, spike' or *acatxar* 'to crouch down', are still lacking a satisfactory explanation, but may be understood in the light of Occitan (Martines 2011; 2012b). In the same vein, territorial forms such as *orandella* 'swallow (bird)' (instead of the more general *oreneta*, *oroneta*, *oronella*) in the Valencian region run parallel to variants explained as words of Occitan origin, which came from the Pallars (Northwestern Catalan).

As stated above, French has a great typological affinity with Catalan and Occitan. French influence has spread throughout the history of Catalan, given the geographic proximity, status and prestige of this language. Again, and particularly due to its annexation to France, Northern Catalonia received the most intense and disturbing influence. The cultural weight of France has been a source for the spread of lexical elements linked with politeness, clothing and gastronomy, into Catalan as well as into many other languages. The present role of English as an international language and for the spread of cultural innovations is well known, but for centuries before that, French was in charge of this task. Basic words such as *xemeneia* or *ximenera* 'chimney', *gec*, *jac*, *jaca*, *jaqueta* 'type of clothing', 'jacket', *pantaló* 'trousers, pants' and so many others originated in French (Barri 1999). During the 19th and 20th centuries, many Valencians and Minorcans emigrated to Algeria. Upon their return, after Algeria's independence from France, they carried with them an interesting set of words of French origin, which have lived on (*tricot* 'sweater', *fregider* 'refrigerator', *cutó* 'knife', *culotes* 'underpants', *pupé* 'puppet').

Italian has also left its lexical mark in many different languages. For instance, cultural terms such as *piano* 'piano', *òpera* 'opera' or *sonet* 'sonnet' are present in Catalan as well as many other languages (Gómez 2012). The most interesting, unique characteristic of the relationship between Italian and Catalan comes from the close territorial relationship between the Italian Peninsula, Sicily and Sardinia and the Crown of Aragon during the Middle and Modern Ages. This meant that Catalan culture became imbued with the Italian cultural milieu, acting as a transmission bridge to the Iberian Peninsula, thereby receiving a large volume of words of Italian origin. Of note were the constant contacts between fishermen and tradesmen in the entire western Mediterranean. Some words of common usage, in some instances even emblematic, such as *orxata* 'almond milk', *sobrassada* 'type of cured meat', *figatell* 'meat processing technique originating in the Valencian region' or *gamatxa* 'type of grape', *colpir* 'to impress', *palesar* (Martines forthcoming), *alatxa* 'type of anchovy' (Veny 2011). This is another field of research that remains open. Regarding Italian influence in the great works of medieval Catalan literature, the novel *Curial e Güelfa* is a very important case, on which recent efforts to describe the traces of Italian have produced interesting results for the diachronic characterization of the Catalan lexicon (Martines 2012a; forthcoming). In stark contrast, in the Catalan spoken in l'Alguer, Italian, a high-status language, has a profoundly influential function; likewise, language con-

tact with Sardinian has injected numerous words of Sardinian origin into Catalan. See Corbera (2000), Bosch i Rodoreda (2012).

As is well known, English has become a kind of international contact language, functioning as a vehicle for the spread of cultural innovations. In a context of intense globalization of the economy and communications, the reach and intensity of English influence, as well as of the culture expressed by means of English, may have effects in the entire world. Hence, the consequences for ethno-diversity have received some scrutiny. It is true that the number of words of English origin in Catalan and most other languages increases exponentially every year. They enter the language through direct contact between speakers, as well as in translation. Advertising, cinema, literature, popular music, fashion, technology, are all thought about, produced and communicated in English. On the other hand, cultural innovations are categorized and expressed in this language: other languages can either choose to adopt them “as is”, subject them to a formal adaptation, or translate them. The margin for creativity is very small (Cabré/Domènech/Estopà 2014). As has been stated in 4, neological and terminological production services (TERMCAT) and normative institutions (Institut d’Estudis Catalans) must often combine an awareness of real social usage, where words of English origin may have already put down roots, with the chance to generate original terms. In Catalan, the process is further complicated by Castilian interference.

Words of English origin may be lexical, with or without graphic/phonetic adaptation: *zàping* (from *zapping*: [ˈzapɪŋ] or [ˈzapɪŋk]), *futbol* (for *football*: [futˈbɔɫ]), *copyright*, *striptease*. They may also be semantic: *audiència* ‘group of people who watch a show’ (from *audience*), or *ratolí* ‘mouse (computer)’ (from *mouse*).

3 Variation and change in the Catalan lexicon

3.1 Diachronic variation and change

Change and permanence, survival and adaptation are two pairs of terms that may summarize the journey of a language (in our case, of Catalan) through time. Indeed, old texts are a testament to the permanence of a significant lexical repertoire in Catalan language from its origins until today. The earliest documents, still written in a language modality halfway between Latin and a nascent form of Catalan (in its written form), already contain words that are entirely frequent today; both basic and more abstract, specialized terms. Hence, it is possible to find the following words in feudal documentation of the 10th and 11th centuries: *aixada* ‘hoe’, *ajudar* ‘to help’, *captar* or *acaptar* ‘to acquire’, ‘to obtain’, *consell* ‘advice’, *dar* and *donar* ‘to give’, *destral* ‘axe’, *engany* (variant of *engan*) ‘deception’, *escudella* ‘bowl’, *finestra* (or *fenestra*) ‘window’, *fogassa* ‘flat loaf’, *home* ‘man’, *manar* ‘to command’, *matí* (or variants *maití*, *maitín*, *matín*) ‘morning’, *menjar* ‘to eat’, *nit* (or *nuit*) ‘night’, *recollir* ‘to collect’, *taula* ‘table’.

On the other hand, parallel to this great lexical trove, maintained from the first written manifestations of the language throughout the centuries until our days, there is a wide range of words in Old Catalan that appear unfamiliar to an average speaker/writer of modern Catalan. Similarly, if the same speaker/writer was shown a sufficiently representative corpus of the evolution of Catalan lexicon, century by century, he or she would surely discover words he or she knows appearing with unfamiliar meanings or in strange contexts. Likewise, he or she would probably notice the lack of certain expressions that are current today and basic to daily life. He or she may even notice words that had been frequent for centuries undergoing a retreat at some point in history, becoming less frequent or acquiring marginal meanings until they are definitively cornered or substituted by other words.

It is the task of diachronic semantics and lexicology to discover the diverse causes of lexical variation and change in language. The fundamental task is to answer questions that only appear simple at first sight: What changed? When did it change? In which usage contexts or places is this change apparent? Among which groups of speakers? How did this change take effect – that is, according to which mechanisms? Which causes, or factors, does this change respond to? Which consequences does the change have in the lexicon of a language? Their aim is to provide a general qualitative and quantitative overview of the degree of maintenance and transformations that the lexicon of a language has undergone. Regarding the Catalan lexicon, this is a project still very much in its initial stages.

A general description of the historical evolution of Catalan has not yet been made. Yet, as stated in 2, there is already an extensive tradition in the study of the possible influences of other languages on Catalan and in the characterization of its vocabulary in the Romance language family context. There are well-known and solid etymological dictionaries, which focus their attention, as is typical, above all on the origin of each word. A few studies follow this path to shed light onto the etymological origin, evolution and territorial coverage of certain words, and a considerable number has focused on editing and studying the language (including the lexicon) of a specific author or text. Furthermore, especially in the past few years, a few studies have applied contemporary linguistic theory to the analysis of semantic change and, particularly, grammaticalization (Cuenca/Massip 2005; Montserrat 2007; Pérez Saldanya 2015; Sánchez López 2015; Antolí 2016; Sentí 2017; Ramos 2016; Garcia Sebastià 2017; Martines 2013; 2015; Martínez 2017; Pons 2018). Nevertheless, there is still a need to go beyond case-studies and consider the general evolution of the lexicon. Taking what is currently known as basic vocabulary as our reference, it may be possible to determine more precisely what has changed and what has remained, and to evaluate the factors that have conditioned lexical-semantic change and the mechanisms of such a change. Our questions may begin to find some answers. This chapter will present a sketch of some of these possibilities.

For entirely methodological reasons, this section will attend to aspects that may affect change in the general lexicon of the Catalan language on the temporal aspect;

in following sections we will focus our descriptions on the lexical configuration processes for different geographical varieties, or dialects; of the different functional varieties, or registers; and of the social varieties, or sociolects, of Catalan.

3.1.1 The same old words, with brand new meanings: semantic change

There is a close relationship between semantic change and cultural context, with all the transformations the latter may experience. This becomes more evident when semantic change is directed – or even directly dictated – by political, economic, religious, scientific and advertising agents. As an illustration, let us consider the change in the definition of the word *matrimoni* ‘matrimony’ after the legalization of same-sex marriage; or the change in the meaning of the word *planeta* ‘planet’, driven by the International Astronomical Union (IUA) in 2006, when it reduced the number of these heavenly bodies in the solar system to merely 8; or the changes in the meaning of *mare* ‘mother’, linked to the advances in reproductive techniques and the evolution in social models of family as an institution.

Although the processes may be determined by numerous factors, everyday communication provides ample evidence of words changing their meanings in accordance with the speakers’/writers’ cognitive and communicative needs. *Correu* ‘[snail] mail’, *adreça* ‘address’ or *bústia* ‘mailbox’ are common, even traditional words, which have very recently expanded their uses. Nowadays, they are more frequently used to designate the basic elements in digital communication technologies: ‘e-mail’, ‘e-mail address’ and ‘inbox’. These metaphors reveal the extent to which new domains are projected over older, better known ones: postal (or ‘snail’) mail, in this case.

Similarly, the current metonymic use of *solter soltera* or its synonym *fadrí fadrina*, a ‘(person) who currently has no romantic partner’ instead of the traditional sense of ‘(person) who has not been married’, may be explained as a reflection of the progressive secularization and transformation of relationships in our society in the past few decades. In the Catalan context, this transformation is represented emblematically by the Spanish Divorce Act of 1981: sexuality has increasingly been liberalized and there is a growing number of romantic partnerships that are not formalized through civil or religious institutions. Currently, the use of *solter soltera* and *fadrí fadrina* provides little information about a person’s marital status; this is a constant in other semantic changes observed within the same domain: *parella* ‘couple’ is increasingly being used to mean ‘person who is in a couple with another’ (*El Carles és la meva parella* ‘Charles is my couple’, i.e. romantic partner): this avoids the use of more (or rather: too) precise expressions related to the gender or kind of union (*marit* or *home* ‘husband’, *muller* or *dona* ‘wife’; *company*, *companya* ‘partner’). It is worth noting that a change in *cultural scripts* usually underlies semantic changes such as these (Wierzbicka 2006).

Apart from semantic changes that undergo meaning *broadening*, instances of the opposite phenomenon, meaning *narrowing*, are not infrequent. Nowadays, it is common to hear *conciliació* ‘conciliation’ used as ‘work-life balance’, *atur* ‘ceasing of activity’ as ‘unemployment’ as well as ‘unemployment benefit’, *mal* ‘ailment’ (“Li han trobat un *mal* al fetge” ‘He has been diagnosed with an *ailment* of the liver’) as ‘cancer’. In our current world, the social and political movements surrounding independence demands in Catalonia have given rise to noteworthy semantic changes: *procés* ‘process’ and *dret a decidir* ‘right to decide’ are found in contexts where they mean ‘social and political movement to obtain Catalan independence’ and ‘demand for a Catalan independence referendum’, respectively.

It may be more accessible, if not easier, to observe and explain contemporary semantic changes, given how a researcher’s own encyclopedic knowledge of his immediate context is often very helpful. But it is often harder to detect and unravel semantic changes that have occurred in the past. The following samples of semantic changes, undergone by Catalan in a variety of periods, illustrate some of the particularities of its lexicon. It seems that both Catalan and Occitan restricted the application of Lat. MŪSCŪLUS ‘muscle’ (diminutive of MŪS ‘mouse’) to ‘shoulder’, *muscle* in both languages. *Blat* (< Cel. BLĀTO) refers in Catalan mainly to ‘wheat, *Triticum*’ (*blat*, in Occitan; *blé*, in French), though in the Middle Ages it was used more broadly to refer to ‘cereal’. For the specific meaning of ‘wheat, *Triticum*’, medieval Catalan used *forment* (< lat. FRŪMENTUM), which is now restricted geographically, or else found in phraseology:

Negun hom estrayn qui aport forment, farina, ordi ne altre *blat* (*Costums de Tortosa*, 1270; Massip 1996, 13)

‘no foreign man (‘from outside the city’) may carry wheat, flour, rye or any other *cereal*’

Civada (< lat. CĪBĀTA, participle of CIBARE ‘to feed’, from CĪBUS ‘foodstuff’) underwent a similar *narrowing* process, from ‘food, especially for beasts’ to ‘cereal used to feed beasts’ and, in the end, to a concrete cereal type: ‘oat, *Avena sativa*’. This is not an unusual process of change, and it puts Catalan once again in the company of Occitan. The Spanish and Portuguese cognate *cebada* followed a metonymic process of the same type, which ended up designating a different grain of the genus *Hordeum* (‘barley’), *ordi* in Catalan and Occitan, *orge* in French and *orzo* in Italian.

Similar *narrowing* processes have been (and still are) very common and have left marks in the Catalan lexicon. That said, generalization or *broadening* cases have been no less frequent. Basic words such as *dona* or *arribar*, now meaning ‘woman’ and ‘to reach a destination’, were initially more restricted in meaning: until well into the Middle Ages *dona*, from lat. DŌMĪNA (fem. of DŌMĪNUS ‘lord’, ‘owner’), maintained a courteous and respectful meaning, opposed to the generic *femna* or *fembra* (< lat. FĒMĪNA). *Arribar*, derived from lat. RĪPA ‘the bank of a stream’, initially meant ‘to reach the shore’ and extended during the Middle Ages to mean ‘to reach a destina-

tion'; this reduced the scope of *venir* 'to come', which had been unchanged for centuries until *arribar* extended to compete with it (Montserrat 2007).

Traditionally, there has also been a focus on semantic changes generated within a *non-denotational* field: those semantic changes that carry a *pejoration* or *amelioration* of meaning; that is, to quote Geeraerts (1997, 99; 2010, 28), a type of "shift towards a (more) negative emotive meaning", in the first case, and "a shift towards a (more) positive emotive meaning", in the second. Examples of a pejorative development can be found in the old *catiu* (< *captiu*) 'captive, prisoner', that became 'unfortunate, miserable'; or *beneit* 'blessed', with a second pejorative meaning as 'dumb' that was attested as early as the 15th century, in addition to the current Catalan meaning as 'good person'. Ameliorative development, on the other hand, can be found in expressions such as *arlot*, which in Old Catalan likely had an important extension in the original sense of 'lowlife (man or youth)', though by the end of the 15th century it is frequently attested, without that negative connotation, to refer merely to young people. In the current Catalan of the Balearic Islands, it is a basic word that means 'boy, girl' (*al-lot al-lota*). Similarly, *canalla* (derived from *ca* 'dog') 'pack of dogs', gave way to 'group of wicked people' and ended up as 'group of children'. In Catalan as well as other Romance languages, *cavaller* developed from its original meaning as 'horseback rider' to 'a man of noble and distinguished character, well-mannered and gallant'.

These labels, *narrowing*, *broadening*, *pejoration* and *amelioration*, are referring to the result of a process of semantic change. In order to explain this process, it is necessary to try to reconstruct the cultural and pragmatic context in which each change may have occurred. Following Grice (1975), with added precisions by Horn (1984), Levinson (1995) and, most recently, Traugott/Dasher (2002), Traugott (2012) specialization or narrowing – which occurs frequently in semantic change – has been explained as a consequence of an *invited inference* triggered in contexts where Relevance-heuristic principles apply. These are connected to the Relevance Maxim ("Be Relevant"): "Narrowing generally involves an R[elation]-based shift from a set denotation to a subset (or member) of that set, representing the salient or stereotypical exemplar of the general category" (Horn 1984, 32). Numerous examples have been provided in the literature, such as the Engl. *deer* (< *deor*), originally 'animal' and later 'any of several ruminants of the family *Cervidae*'; or *corn*, used to mean 'whatever grain is the most important cereal crop of a particular region, e. g. *wheat* in England, *oats* in Scotland, *maize* in Australia or the New World' (Horn 1984, 32–33). *Broadening* is also an R-based change, involving "the generalization of a species to cover the encompassing genus to phylum, from subset to superset" (Horn 1984, 35). Examples are Lat. *pecúnia* 'property, wealth in cattle', which became 'money', as well as the generalization of commercial brands to refer to concrete products: *xerox*, *kleenex*, *thermos*, etc.

The R-heuristic principle involves the speaker/writer using "an expression that is less explicit than it might be", which triggers the *invited inference* that enables the addressee/reader to reconstruct the full interpretation of the expression which, once

semanticized, will end up codifying a new meaning (Geeraerts 2010, 231; Martines 2015). This occurs with the support of a concrete cultural context: notice how both the current and old samples of *narrowing* and *broadening* we provide reflect cultural developments, interpretations and conceptualizations: legal developments, change in scientific knowledge, views on family, technological innovations, romantic relationships and associated values, the human body, human and animal feeding habits, the image of women in medieval society, the importance of navigation in the relevant society, etc.

A regular semantic change pattern in the provided examples is the tendency towards *subjectification* of meaning, according to which “meanings tend to become increasingly based in the speaker’s subjective belief state/attitude toward the proposition” (Traugott 1989, 34; 2012, 557). In a few of these cases, there may even be a further tendency towards *intersubjectification*, that is, the semantization of the speaker’s/writer’s attitude towards the addressee/reader, their knowledge of the world, opinions, self-image, social status or courtesy (their *face*). This element has an impact on changes such as *blat* ‘cereal’ > ‘wheat’ (plausible in a context where both speaker/writer and addressee/reader have a shared prototype of *Triticum*), but also instances more closely linked with attitudes and self-image (cf. supra *solter soltera*, *fadrí fadrina*; *dona*). This is even clearer in instances of *pejoration* and *amelioration*, with a “physical > intellectual > evaluative” evolution; consider Dutch *dom* ‘dumb’, that evolved from ‘unable to speak’ to ‘unintelligent’ and, more recently in Belgium, “has come to express a negative attitude on the part of the speaker, meaning ‘annoying, cursed, bloody’” (De Smet/Verstraete 2006, 374).

Underlying all the examples of semantic change presented thus far, metaphor and metonymy function as fundamental cognitive mechanisms to understand the world. The former is based on the projection of a source conceptual domain onto a target conceptual domain. This projection is founded upon the similarities between domains, as judged by the speaker/writer-conceptualizer and other speakers. The latter is based on contiguity (spatial, temporal, and cultural) between concepts. Metaphor and metonymy are essential tools to ensure the permanent adaptation of the lexicon to the speaker’s/writer’s-addressee’s/reader’s cognitive and communicative needs. Ultimately, these tools enable both language creativity and continuity. Words that have developed characteristic meanings in Catalan possess metonymic and metaphoric origins, such as *carena* (< lat. CARĪNA), a nautical term meaning ‘keel of a ship’, which has developed metaphorically to mean ‘portion of a body that elevates lengthways higher than the surface it belongs to’, and, more concretely ‘mountain peak’ and ‘ridge of a roof’. *Enyorar* ‘to miss’, which evolved from lat. ĪGNŌRĀRE ‘not to know’, underwent a metonymic (cause-effect) transition into Spanish (*añorar*): it expresses the feeling experienced by someone who has not heard any news from someone or something dear to them. Also related are *panteixar* or *pantai-xar* ‘to pant, to gasp’, which originated in Lat. PANTASIARE (popular variant to PHANTASIARE) ‘dream’ and ‘have a nightmare’, etc.

Cultural developments tend to put people's cognitive strategies to the test when encountering unfamiliar objects. The speaker/writer is often confronted with the question: "What is this thing?" Metaphor and metonymy are useful resources to tame the strangeness, and they have been useful through the history of Catalan lexicon. As a small sample, let us look at some semantic changes that likely began in the Middle Ages, just as Catalan began the minoritization process that has characterized part of its history. The discovery of America brought to the Old World a host of animals and plants that required new names. In a few cases, just as in other languages, Catalan borrowed the indigenous terms, which arrived in Europe with these objects, though slightly altered through Spanish contact. Such was the case of *patata* 'potato', adapted from Sp. *patata* (presumably a blend of *papa* 'potato' and *batata* 'sweet potato'); or *tomàquet* (and its variants: *tomaca*, *-ta*, *tomàtiga*, etc.), based on the sp. *tomate* 'tomato' (itself based on Aztec *tomatl*). In other cases, similarities were found with existing Old-World products, which influenced their labeling: *atzavara*, of Arabic origin, referring to aloe, became *Agave americana*. *Pebre* (< lat. *PĪPĒR*), meant until the 16th–17th century 'pepper', after which it was used to refer to vegetables of the *Capsicum* genus, in various derived forms (*pebre*, *pebrot*, *pebrera*, etc.). In addition to this, certain varieties received metaphoric (*bajoca* 'pod', *coral(et)* 'coral', *vit* 'penis', *ditet* 'little finger') and metonymic (*pesteta* 'stench' > 'spice') terms (Martines 2000a). Similarly, *Zea Mays*, the quintessential American cereal, was metaphorically integrated into Old-World vegetable categories: *blat* (*de moro*, *morenc*, *de les Índies*), *dacsa*, *panís*, *milloc* 'moor wheat, wheat of the Indies', *dacsa* (most likely a cereal like spelt), *panís* and *mill* (or *milloc*) '*Seratia italica*'. The American wild turkey (*Meleagris gallopavo*) received metonymic denominations based on its origin: *gall dindi*, *gall de les Índies*, *indià* o *indiot*, etc. 'rooster of the Indies', 'indian rooster', or else based on its song: *pioc*, *titot*, *tito*, etc.

3.1.2 Forgotten words

As illustrated above, the history of the lexicon of any language is a chance to observe new meanings emerge. These new meanings may coexist with previous meaning (which leads to polysemization) or come to substitute them. Along these lines, another fundamental manifestation of lexical change is the disappearance of words that may have been frequent in previous epochs. Words becoming less usual, that is, restricting their meanings to concrete contexts (phraseology, onomastic, argot, regional varieties) or even ceasing to be used altogether and becoming archaisms, are phenomena determined by pragmatic and cultural factors like those suggested above for semantic change; let us briefly consider some samples.

Inevitably, changes in material culture, conceptual systems or cultural scripts in a society are some of the most important motivations: objects or concepts may disappear and, consequently, the words that designated them cease to be used or start

being perceived as old-fashioned or even strange. Let us think of the more or less basic elements that have disappeared from daily life: *alcandora* and *gonella* ‘types of chemise’ were medieval undergarments worn by men and women (in some regions, the latter existed until the 19th century); *quarteró*, *almut* and *jovada* are measuring units for weight/capacity and surface area, popular before the metric system became dominant; traditional trades that were common up until recently have become less frequent, and hence, their names may remain in the memory of the speakers (especially of older generations’) but are no longer used in daily speech: *bugadera* ‘woman who washes other people’s clothes’, *dida* ‘woman who nurses other peoples’ infants’, *esmolet*, *esmolador* and *esmolaire* ‘person who sharpens knives’, *matalasser* or *matalafer* ‘person who makes or fixes mattresses’. The death of an *albat* ‘child who died before reaching the age of reason’, used to be, up until the 19th and early 20th century, a matter not mourned excessively, and sometimes even celebrated with song and dance: it was believed that the innocent soul of the *albat* had gone straight to Heaven. Nowadays, both the custom and the word persist only in the memories of elderly speakers.

Frequently, words disappear from daily use because they are substituted by others. Here, as in the emergence of new meanings, the influence of science, economics, religion and folklore are not rare. As we shall see when we discuss functional variations, under the influence of scientific argot, *lleu* and *empelt*, *empeltar* gave way to *pulmó* ‘lung’ (15th–16th centuries), and *vacuna* ‘vaccine’ and *vacunar* ‘to vaccinate’ (19th and 20th centuries), respectively; it seems that *budell* and *endanyar* or *endenyar* have likewise been displaced by *intestí* ‘intestine’ and *infectar* ‘to infect’. In our contemporary society, where old age and the end of life are taboo subjects, *vell vella* and *ancià anciana* are best avoided when speaking about ‘people of advanced age’ and are often replaced by *gran* or *major* ‘older’, *madur madura* ‘mature’, or (*de la*) *tercera edat* ‘(of the) third age’. Similarly, the term *volp*, from Lat. VŪLPES, to mean ‘fox’, an unpopular, even feared animal, was replaced by *guineu*, *guilla* or *rabosa* in the Middle Ages.

Words’ roles as sociolinguistic and pragmatic markers are decisive in determining their usage and survival. Indeed, any linguistic element may perform this function, but vocabulary is probably the field that does this most tangibly. Take, for instance, the factors linked to politeness in a context of intersubjectivation. An illustrative case is the retreat, since the end of the 15th century, of the verb *amar* ‘to love’ in the face of the learned word *estimar*; this verb, which originally meant ‘to determine or estimate the extrinsic (money) value of a thing’, had already developed further meanings such as ‘evaluate’ and merged the medieval scientific and psychological terminology used to explain perception (*vis aestimativa*). As we have shown in Martines (2015), the use of *estimar* for *amar* afforded an unmarked term for the Lat. *amor hereos*, free of the negative connotations that had been accrued by *amar*, within a cultural context that viewed this type of love as a true pathology. The earliest attestations of this lexical particularity in Catalan show how *estimar* is present in a dialogue where one of the

characters tries to make his interlocutor believe that he ‘loves honestly’ and is not under the spell of that sort of pathological love (Joan Roís de Corella, *La istòria de Leànder y Hero*, 15th century; Martos 2001). In contrast to what has previously been asserted, this change originated in formal registers and later extended into everyday language.

There is a large set of words in the basic vocabulary of Catalan that have been totally or partially substituted, after an extended period of coexistence with Castilian. Throughout the Middle Ages, *emblar* gave way to *robar*; *toldre* to *llevar* ‘take away’; *llaguiar* to *tardar* or *trigar* ‘take, delay’; *membrar* to *recordar* ‘to remember’. Indeed, the causes of changes such as these are varied, and perhaps not entirely clear yet. In cases such as *frare* and *sor* replaced by *germà* ‘brother’ and *germana* ‘sister’, it seems due to an approximation to the Spanish language (*hermano hermana*) (Colón 1976), a factor that has gained in importance since the Modern Age. A detailed corpus analysis of other cases shows syntactic and semantic conditions that should not be taken for granted. For instance, the expansion of *arribar* ‘to reach the shore’ vs. *venir* ‘to reach a destination’ was linked to cultural factors (the cultural significance of navigation) as well as linguistic ones: *arribar* made it possible for the focus to be placed on the goal of the process or movement (in periphrastic constructions), while the image scheme of *venir* also integrated origin and the notion of approaching the goal (Montserrat 2007). In the same vein, the use of *metre* ‘to put, to place’ as a lexical and *light verb* (in collocations and periphrases) decreased significantly in Catalan since the 15th century, until it remained only in derived vocabulary (*permetre* ‘to permit’, *prometre* ‘to promise’, etc.); it was replaced by *posar*, a verb that initially presented fewer semantic and constructional restrictions than *metre* (Pons 2018). In contrast, Lat. *mittere* cognates are frequent in Spanish (*meter*) and in other Romance languages (Fr. *mettre*, It. *mettere*).

Traditionally, homonymy between two words, especially basic and frequent items, has been viewed as an instance of lexical change: homonymy may interfere in communication, and difficulties may be eliminated through substitution of one of the terms. A classic example is provided by Gilliéron/Roques (1912): Late Lat. CATTUS ‘gato’ and Lat. GALLUS ‘gallo’, given the evolution of /l:/ > [t], characteristic of Gascon, converged in *gat* in this variety of Occitan. Homonymy was resolved through the substitution of the latter by derivations of Lat. PŪLLUS (‘young animal’), by *hasan* ‘pheasant’ or *veguièr* (< Lat. VĪCĀRĪUS ‘curate’). As stated above, explanations based on *homonymiphobia* (Geeraerts 1997, 130) have limited explanatory power, as they seem to neglect part of the context of communication (Traugott/Dasher 2002, 53). On the other hand, although homonymy certainly decreases iconicity and violates the principle of isomorphism (“one-form: one-meaning”), which is linked to communicative efficiency, it is worth remembering that not all cases of homonymy are detrimental to communication. This violation is similar for polysemy, a phenomenon that is entirely frequent and even constitutive of natural language. In any case, it seems likely that the retreat of the verb *llavar* ‘to wash’ in Eastern Catalan (a dialect in which

unstressed vowels /a/, /e/ and /e/ are neutralized in [ə]) and substitution by *rentar* ‘to rinse’ can be explained through the collision of *llavar* and *llevar* ‘to remove’, only a homonym in this particular dialect.

Pragmatic marking is probably linked to the substitution of words such as *sutze* or *sützeu* ‘dirty’ by *brut* (initially ‘characteristic of beasts, brutal, irrational’); of *calcigar* ‘to tread’ by *trepitjar* (initially, ‘grape-stomping to make wine’) and, later, depending on the region, by *xafar* (initially, ‘to smash’); *esquinçar* ‘to tear fibrous matter’ by *estripar* (initially, ‘to gut’), *esguellar* (or *-allar*) (initially, ‘to rip out a branch’). These are cases in which words of a more expressive origin, or more intense value, have gained ground. This tendency is not rare in lexical evolution: an onomasiological study may reveal that it seems to affect the expression of certain concepts more than others. For instance, [CHILD] has received various pragmatically marked denominations through time, which, depending on the region, have forced older terms to retreat into formal, territorial or literary domains: (*infant*, *minyó*, *nin*): *noi noia* (a reduced and expressive variant of a diminutive *ninot* of *nin nina* ‘child’, old Eastern Catalan), *marrec* (‘lamb’: Eastern Catalan), *al·lot al·lota* (see above; Balearic Islands), *fillet filleta* (diminutive of *fill filla* ‘son daughter’; Balearic Islands), *boix* (originally, *Buxus sempervirens* ‘shrub’ and ‘object elaborated with its wood’; Eivissa).

Language contact has been (and still is) a point of entry for new words, loan words as well as interference. A diachronic description of this can be found in 2, and we will return to it in 4, when we deal with the current situation of the Catalan lexicon in a context still greatly influenced by Spanish (in Catalan-speaking domains that are situated in the Spanish territory), French (in Northern Catalonia) and Italian (in the city of l’Alguer, in the Italian island of Sardinia). All of this will be considered beneath the umbrella of globalization.

Contact between territorial registers and varieties (or dialects) of the same language is a significant source for the propagation of some words, and the subsequent substitution of others. Lexical exchange between formal and colloquial registers has remained constant; both are often linked to a specific social or professional class, like the infiltration of scientific terms into everyday vocabulary, to which we have alluded above. This phenomenon, both old and current, will be further discussed in our section on functional variation (3.3). Here, we will also briefly address the spread of standard languages and their influence on these processes of lexical renewal.

It is not rare for words that have progressively disappeared from everyday use – as outlined above, due to the disappearance of their referent, or because they have been replaced by other words – to find shelter, for instance a) in phraseological units of whose original meaning the average speaker may not be aware: *lleu*, now almost replaced by *pulmó* (‘lung’), remains above all in idioms (*Treure el lleu a algú* ‘to cause someone to make a great effort’; *Més content que gat amb un lleu* ‘very satisfied’); *censal* or *vectigal* were terms that originated, respectively, in the financial and tax system of past eras; they appear now in phrasal units such as *costar un censal* ‘to be

very expensive' or *donar vectigal* 'to cause trouble'; b) in specialized language: *metre* (cited above) means 'to put a net back into a ship' (sailing terminology); c) in onomastics: the preposition *jus* 'below' and its derived adjective *jussà jussana* 'inferior', in decline since the late Middle Ages, remain, for instance, in *Casadejús*, a last name meaning 'house down below', or in the name of the region *Pallars Jussà*.

As we shall see in 3.2, words that may appear archaic in standard language do occur, sometimes quite frequently, in other dialects or geographical varieties.

3.1.3 New words

The creation of new words is the third form of lexical change we will address here, although only briefly. In general, it is a function of cultural, pragmatic and cognitive motivations, which trigger – as shown above – changes in meaning or the decline of specific words. We will not elaborate on the morphological mechanisms of lexical creation in Catalan, which have been explained in 75.1 Word Classes, Inflectional Categories and Paradigms in this volume. Let us briefly present them from a diachronic perspective.

Derivation (prefixing and suffixing) and compounding are the main mechanisms of lexical creation in Catalan, from its very beginnings. Though at times it may be difficult to distinguish whether a given word was directly inherited from Latin or generated later in Catalan, still, the earliest texts show derivation and affixes which are still productive nowadays, as well as words composed according to patterns that are still viable: here are some old examples from the 10th and 11th centuries, in chronological order.

a) i. Suffixing:

Guardiola 'little guard's post' (year 983: already a toponym: [[*guàrdia* 'a guard's post']_N *ol* (a)_{elN}], *pujol* 'little hill' (1076: [[*puig* 'hill']_N *ol*]_N);
pellissó 'fur garment' (in the half-Latinized Romance form *pellicone*; 986: [[[*pell* 'skin, leather']_N *ís*]_N *ó*]_N);
podadora 'pruning-knife' (986: [[*podar* 'to prune']_V *ador* (a)_{elN}];
guerrear 'to wage war' (variant *guerriar*; 1055–1089: [[*guerra* 'war']_N *ejar*]_V),
llancejar 'to spear' (12th century; [[*llança* 'spear']_N *ejar*]_V), *pecejar* 'to break, to destroy' (variant *pecijar*; 12th century: [[*peça* 'piece']_N *ejar*]_V), *pledejar* 'to litigate' (12th century: [[*plet* 'lawsuit']_N *ejar*]_V);
albereda 'poplar grove' (1076: [[*âlber* 'poplar']_N *ed* (a)_{elN}];
casella 'little house' (1076: [[*casa* 'house']_N *ell* (a)_{elN}];
cavaller 'knight' (1080: [[*cavall* 'horse']_N *er*]_N);
establiment 'permission to possess armed forces' (1080: [[*establir* 'to establish']_V *ment*]_N);
valença 'protection' (1080: [[*valer* 'to be worthwhile']_V *ença*]_N);

rancurar ‘to accuse’, ‘to claim something before someone’ (1115: [[*rancura* ‘grievance’]_N ϕ]_V);
franquesa ‘exemption’ (variant in *-esa* and *-ea*; 12th century: [[[*franc*]_{ADJ} ‘exempt’]*esa/ea*]_N);
oradura ‘madness’ (12th century: [[*orat*]_{ADJ} *ura*]_N);
rauberia ‘depredation’ (12th century: [[*robar*, *raub-* ‘to pillage’]_V *eria*]_N);
saber ‘wisdom’ (12th century: [[*saber*]_V ϕ]_N), *plaer* (variant *plader* (*pla*[z]*er*); 12th century: [[*plaer*]_V ϕ]_N); *haver* ‘property, goods, money’ (12th century: [[*haver*]_V ϕ]_N);
verament ‘truly’ (12th century: [[*ver* (*a*)_f]_{ADJ} *ment*]_{ADV})...

ii. Prefixing:

sobrepellís ‘surplus’ (in the half-Latinized Romance form *subrepelliceo*; 986: [*so-bre* [[*pell*]_N *ís*]_N]_N);
recollir ‘to collect’ (1036; [*re* [*collir* ‘take’]_V]_V);
acollir ‘to accept’ (1080: [*a* [*collir* ‘take’]_V]_V), *afollar* ‘to harm, to destroy’ (1080; [*a* [*follar* ‘to hit; to stomp on something’]_V]_V);
esbalçar ‘to topple, to destroy a building’ (1085: [*es* [*balç* ‘precipice’]_N]_V);
contradir ‘to act against, to oppose someone’ (12th century: [*contra* [*dir* ‘say’]_V]_V);
desfer ‘to destroy’ (12th century: [*des* [*fer* ‘to do’, ‘to make’]_V]_V), *sobreviure* ‘to survive’ (12th century: [*sobre* [*viure* ‘to live’]_V]_V);
traspasar ‘to transgress’ (variant *trespassar*, 12th century: [*tras* [*pasar* ‘to pass’]_V]_V)...

b) Compounds:

franc alou ‘territorial property exempt of any feudal charge or claim’ (1085: [[*franc* ‘libre’]_{ADJ} + [*alou* ‘allodium’]_N]_N);
manllevar ‘to borrow’ (11th–12th century: [[*man* ‘hand’]_N + [*llevar* ‘to raise’]_V]_V);
cansalada ‘bacon’ (< ‘meat preserved in salt’; 1189: [[*car* ‘meat’]_N + [*salada* ‘salted’ (past participle of the verb *salar* ‘to preserve in salt’)]_V]_N);
afer ‘matter’ (12th century: [[*a* ‘to’]_{PREP} [*fer* ‘to do’, ‘to make’]_V]_N);
menysprear ‘to despise’ (12th century: [[*menys* ‘less’]_{ADV} + [*prear* ‘to put a price on something’]_V]_V);

It is important to remember that these morphological procedures rely on cognitive operations. For instance, derivation products such as the above are not merely mechanic, morphological operations of suffixing or prefixing: in *traspasar* and *sobreviure* (still current words) there is also a metaphoric projection of movement schemata ([TO GO/LIVE BEYOND A LIMIT]) to express the concept of [TO DISREGARD A LAW] and [TO STAY ALIVE PAST A DETERMINED TEMPORAL LIMIT]. In old Cat. *rancurar*, a cause-effect metonymy has played a role: [TO ACCUSE] is the consequence of the [RESENTMENT] (*rancura*) that someone’s actions have caused on someone else.

Similarly, for *esbalçar* (still current, though territorially), [PRECIPICE] is the departure point to generate a metonymic projection: the effect of tossing something down a precipice (*balç*) is its destruction (cf. *timba* ‘precipice’ > *estimbar* ‘to throw down a precipice’). Note how, in all cases, subjectification is a constant.

A few details are relevant from a historical perspective: a) the progressive retreat of affixes that used to be more ‘alive’; such is the case with diminutive suffix *-ell* (before *-et*); b) contrast with the expansion of others, initially from a formal domain, like adverb-generating *-ment* (as is well known, as a result of the grammaticalization of the noun, *ment* ‘mind’) or, later on, the adjective superlative *-íssim*; c) the introduction of affixes from other languages; such is the case of Occitan *-aire*, already present in the names of medieval trades with a verbal base (*paraire*) and progressively extended also to nominal bases, with other values; d) territorial differentiation in the use of certain affixes: variant *-ea* for *-esa* (*bellea/bellesa* ‘beauty’) has become more characteristic of the Catalan spoken in the Valencian region; the suffix *-era* ‘desire for [BASE CONTENT]’ (*ballera* ‘desire to dance’) and the variant *-idat* (for *-itat*: *velocidat* (for *velocitat*) ‘speed’) are, respectively, more frequent, and characteristic of the Balearic Islands.

As Catalan entered the writing age, throughout the Middle Ages and the Modern Age, especially when it began producing specialized literature (in theology, astronomy, law, medicine, veterinary medicine, mathematics), these mechanisms developed further, together with the introduction of learned word constituents. Ramon Lull is often cited as one of the first writers to deliberately employ derivation in the generation of Romance terminology. Beyond this, we can observe a rapid emergence of scientific terminology from the 13th century onwards. This process, analogous to that of other languages, both within the Romance context and outside, has been sustained until our times, even intensifying as communicative and cognitive needs have become more diverse (due to the access of Catalan to formal registers). One only needs to recall the introduction of Greek and Latin terms, with varying degrees of adaptation, or the creation of new ones based on their constituents. A portion of this lexical repertoire, while initially linked to formal registers, has been entering the everyday vocabulary, sometimes even replacing its traditional equivalents (cf. 3.3).

As we have shown in 2, and as we will discuss in the context of contemporary Catalan in 4, contact with other cultures and languages has been a permanent pathway for the entry of new concepts and words into Catalan. We have illustrated the main points of entry for loan words in the lexicon: commerce, migration, translation and, more recently, traditional and social media and other means of sharing knowledge and information. The sociolinguistic context is also a key influence on the status of Catalan: in a state of minoritization, loans are experienced as interference.

3.2 Dialectal variation

The Catalan language is not particularly marked by dialectal variation. The main dialect blocks (Eastern and Western Catalan) are usually distinguished by means of a phonetic factor: reduction or neutralization of /a/, /ε/ and /e/ to [ə] in unstressed position (Milá y Fontanals 1861, 461; Veny ⁴1983, 14). The remaining traits, phonetic as well as grammatical and lexical, are neither as systematic nor as clearly distinctive, in particular the latter two. From a quantitative perspective, only a small part of the Catalan lexicon presents significant dialectal variation, and it would be fair to say that this variation does not go beyond the one observed in other, fully normalized languages, like Spanish in Spain or English in Britain. A representative (open) list of words that illustrate lexical variation between Western (W) and Eastern (E) Catalan is available in Veny/Massanell (2015, 102). From this list, the following words approximate most closely the limits defined by the trait [presence/absence of vowel reduction] as described above: *mirall* (E)/*espill* (W) ‘mirror’ (ALDC, vol. 2, map 2), *llombrígol* (E)/*melic* (W) ‘navel’ (ALDC, vol. 1, map 61) or *romaní* (E)/*romer* (W) ‘rosemary’ (ALDC, vol. 5, map 1099).

Other words in this list, though corresponding less precisely with the classic twofold distinction, do serve to mark a boundary between other territorial units, and even to establish coincidences between sub-varieties within different blocs. The samples we will consider next may well be diverse typological manifestations of lexical-semantic variation.

In any given language, one concept may be expressed with different terms in different territories (*geosynonyms*). This variation may be highly local or regional: fruits, vegetables, birds, fish, insects, are all domains in which this phenomenon occurs very frequently, in many languages. Here, we will focus on the more basic, everyday vocabulary. For [CHILD (UP TO 8 YEARS)], the main geosynonyms are: *xiquet* -a, in the Valencian region, south (S) and west (W) of Catalonia and most of Catalan-speaking Aragon (that is, most of Western Catalan) and spots in South-western (SW) Cat.; *nen* -a, in Central Eastern (CE) Cat. and the southernmost part of Western Cat.; *nin* -a, in Majorca and Northeastern (NE) Cat., both situated phonetically within Eastern Cat.; or *noi* -a, the term for [BOY/GIRL] in Central Eastern Cat., which in the same language may be extended at times to the younger child, or its diminutive *noiet* -a. The concepts [TO SWEEP] and [BROOM] display an interesting distribution: *agranar* and *granera* are found in the Balearic Islands (phonetically within Eastern Cat.) and most of Western Cat. (the Valencian region, Catalan-speaking Aragon, western and southern Catalonia) and areas of Northern (N) Cat. (phonetically also within Eastern Cat.); *escombrar* and *escombra* occur in CE Cat. and in part of Northwestern (NW) Cat. In any case, *escombrar* appears with different but related senses of [TO SWEEP]. For instance, in some areas of the Valencian region, it may mean ‘to sweep the bread oven’ (also in the Balearic Islands), ‘clear one’s throat’, ‘to let almond blossoms fall’, ‘to give trees a trim’ (Martines 2018). Indeed, there are words that can identify concrete

territories: *aplegar* ‘to arrive’, ‘to reach’ is characteristic of the Valencian region, where it coexists with the more general *arribar*; *al·lot -a* ‘child’, already cited above, it is characteristic of the Balearic Islands, and *pallago -a* is characteristic of NE Cat.; *eixaviuro* replaces *esternut* ‘sneeze’, in CE Cat.; *padellàs*, for *test* ‘bucket’ and ‘piece of a broken bucket’, occurs in NW Cat. [MEASLES] has some characteristic labels: *sarampió* NW Cat. (W of Catalonia, Catalan-speaking Aragon, N of the Valencian region), *xarampió* in Eastern Cat. (N and C), *rosa* (Majorca, Minorca, some northern areas of the Valencian region, and some areas of NW and E Central Cat.) and *pallola* (most of the Valencian region and the island of Eivissa). *Eixir* ‘to exit’ replaces *sortir* in most of the Valencian region, parts of Catalan-speaking Aragon and of EC Cat. (Northern Catalonia).

Word semantics may also present territorial variation. The Balearic Islands and the Valencian region have maintained the distinction between *banyar* ‘to wet’ (*Agafa el paraigua si no vols banyar-te: vol ploure* ‘Bring your umbrella if you don’t want to get wet: it’s about to rain’) and *mullar* ‘to dunk, to dip, to soak’ (*M’agrada mullar les galetes en la llet* ‘I like to dunk/soak my cookies in milk’); CE Cat., uses *mullar* for both senses. Speakers in the Valencian region use *llavar* meaning ‘to wash’ (*He llavat la roba amb sabó* ‘I washed my clothes with soap’) and *rentar* as ‘to rinse’ (*M’he rentat les mans a l’aixeta* ‘I rinsed my hands on the tap’); the remaining varieties use *rentar* in both cases and do not know *llavar*.

If we consider current instances of lexical-semantic variation from a diachronic perspective, we may find that:

- a) the current territorial distribution for certain words or meanings may not coincide with past stages: contemporary lexical isoglosses have been configured over long periods of time. The previously discussed example *nin -a* (attested for the 13th century) was most likely in general usage in medieval times; *xic -a* and its diminutive *xiquet -a* must have spread especially in Western Cat. and, above all, in the south since the end of the Middle Ages: this is the usage of the adjective *xic -a* as a noun, which my research has attested for the 13th century, and whose usage was generalized in the Middle Ages; *noi -a* is the expressive alteration of diminutive *ninai -a* (< *nin -a*), documented in the 17th century; *nen -a* and *pallago -a* are latecomers (18th and 19th centuries, respectively): both stand in continuity with equivalent forms in Castilian and Occitan (PALDC, map 306). We have attested *granera* and *agranar* in 15th and 16th century texts in CE Cat. At that time, *escombrar* had the more generic meaning ‘to clean’, ‘to remove clutter from a place’; this generic use connects the meanings of *escombrar* in territories where *agranar* is used to mean ‘to sweep’ (Martines 2018). *Sarampió* and *rosa* are already medieval (14th and 15th centuries), *pallola* came later (16th century) and it stands in continuity with similar forms in Aragonese and southeastern Castilian; it is worth noting that, in places where *sarampió* (or *xa-*) and *pallola* are used to mean ‘measles’, *rosa* may refer to other skin conditions (Martines 2002). *Eixir* and *sortir* used to have different general extensions and different semantic nuances, which

were, up until very recently (18th–19th centuries), partly preserved: the first was the generic ‘to exit’; the second expressed the idea of ‘storming out impetuously.’ *Banyar/mullar*, in the senses mentioned above, and *llavar* ‘to wash’ appear as well in medieval texts attributed to dialects that no longer use it (Martines 2012a).

In contrast, some of the cases cited must have originated as early as the Middle Ages: *aplegar* ‘to arrive’, ‘to reach’, now characteristic of Valencian Cat., is found mainly in medieval texts of this region.

- b) As illustrated in some of the examples above, the survival of words and meanings in peripheral/discontinuous areas, and their absence in central regions, may be an indication that such words and meanings have been generalized in previous eras, and that they must have experienced withdrawal in the face of innovations, which often irradiate from urban centers. What was general once, in a given stage of the evolution of the language, may become marginal: often these usages are described as archaic.
- c) Traits currently functioning as lexical-semantic markers seem to fade into the past, the further one looks back. This is still an open question, as the progressive historical configuration of the lexicon in Catalan dialects has not been definitively established yet. The chronology of the origin and spread of phonetic and morphological units is much better established due to the more systematic character these traits as dialectal markings (Duarte/Alzina 1984), quite unlike the lexical-semantic domain. As we have explained in Martines (2012a), a selection of lexical-semantic traits to distinguish the diverse Catalan dialects diachronically is still missing, a selection which should be based on quantitative data, qualitative analysis, and text corpora. The guiding question would be: *What are the lexical-semantic markers that have characterized Catalan dialects through time?* The results could be applied to the description of the mechanisms, factors, chronology and areas of propagation of lexical-semantic change associated with dialect variation. It may also allow us to improve the description and understanding of old texts, their geographic attribution, or even allow us to define more precisely the role each dialect variety may have played in the configuration of the standard language or written medieval koine, and its progressive territorial diversification. This way, it has been said, for instance, that Valencian Catalan played an important part in developing the standard model, which was used in literary, religious and scientific texts, as well as documentation of the Royal Chancery; in what pertains to vocabulary, it has been considered in awareness of the fact that some of the lexical options in the model seem to be characteristic of (current) Valencian Catalan. One good example of this is found in the speaker’s judgements found in the *Regles d’esquivar vocables o mots grossers o pagesívols* (‘Rules to avoid rude or peasant expressions or words’). These early (late 15th century) rules of linguistic correction give recommendations on words and variants that may already have been characteristic of this variety of Catalan by the end of the Middle Ages (Badia 1999; Colón/Ferrando 2011; Ferrando 2011). In the same vein,

although some expressions have historically been seen to belong to Central and Southern Catalan, and to be foreign to the northern varieties, historic documents and traditional usage indicate that they were also present in northern regions (for example *fleca* ‘bakery’, *tempir* ‘prepared soil’ or *trespol* ‘mortar’; Martines 2012c); or, as we have discussed above, vocabulary proper of the periphery has been found in the entire Catalan domain in the (sometimes not even very distant) past.

The weakening of a formal reference language model, due to political factors since the Middle Ages (see 711 The Growth and Expansion of Catalan (1213–1516)) is likely to have strongly favored the dialectal diversification of the Catalan lexicon. This factor interacted with two further fundamental cultural and sociolinguistic elements:

- a) The emergence of partially different social, political and economic structures in each region of the Catalan-speaking domains. These structures define the cultural context in which the cognitive and communicative needs of the speakers emerge and change. In a situation where the channels that ensure the generalization of lexical innovations throughout an entire language domain (that is, through the ages, the administration, religious discourse, theater, popular and cultivated literature, teaching, media, etc.), the kinds of lexical-semantic change processes that may be expected in language changes, such as those described in 2.1, may intensify centrifugal language change tendencies. Without these channels, the center(s) and periphery(-ies), always in a fluctuating balance, may take divergent roads: the center may generate changes that may not reach the peripheries, which will retain archaic solutions; or else the periphery may generate changes that will not become generalized or reach the center. A large part of the examples presented above are reflections of these dynamics, and they may affect both language categorization and expression of cultural novelty, and the production of new terms for known concepts. We will refer to this issue once more in 4.
- b) The historical coexistence of Catalan with other languages in each territory. Western Catalan has historically been in contact with Aragonese due to geographic proximity, and speakers of Aragonese have been present within the Catalan majority since the early Middle Ages in the Kingdom of Valencia: this may explain the presence of Aragonese words in this region, or of words that stand in continuity with that language domain (*corder* ‘lamb’ (14th century), for *anyell*; *carrasca* ‘holm oak’ (17th century) for *alzina*); intensified contact with Arabic in the west and south of Catalonia, Valencia and the Balearic Islands, explains the more extended presence of Arabic terms in those areas (*farnaca* ‘leveret’, for *llebretó* (or *-ató*); *aladroc* ‘anchovy’, for *seitó*; *almàssera* or *tafona* ‘olive mill’, for *molí d’oli*); contact with Occitan and French in northern Catalonia has left some words of Occitan origin (*feda* ‘sheep’, for *ovella*; *veire* ‘cup’, for *got*) and French origin (*muleta* ‘scrambled eggs’, for *truita*; *canart* ‘duck’, for *ànec*) that were unusual in other regions; in l’Alguer (Sardinia, Italy) contact with Sardinian and

Italian has left numerous words of Sardinian origin (*anca* 'leg', for *cama*; *eba* 'mare', for *egua*) as well as of Italian origin (*indiriz* 'address', for *adreça*; *sécol* 'century', for *segle*) (Veny 1983, 102).

The effects of Castilian interference have not spread throughout the Catalan domain in a homogenous fashion. The more intense effects have been registered, for instance, in urban areas with a higher historical concentration of immigrant populations, or where the higher social classes have been more open to cultural innovations that arrived by means of Spanish, or even preferred to speak in this language. Likewise, language contact in border regions has been more intense, as these areas have historically experienced contact as a fact of daily life. Words of Castilian origin have ended up being characteristic of various Catalan varieties; even if we only focus on words that have become part of the normative repertoire, we find words such as *armilla* 'waistcoat' (18th century) and *vano* 'fan' (17th century), borrowed from *almilla* and *abano* (or *abanico*), which was highly unusual in the Balearic Islands and the Valencian region, where *jupetí* and *ventall* survive instead; *xop* 'black poplar' is used in the Valencian region, as well as in most of Western Catalan, and it is already present in 14th century Valencian texts (probably thanks to the logging industry which transported this variety of timber from Castile). Conversely: *cercar* 'to search for' is now characteristic of the Balearic Islands, whereas *buscar* 'to search for' (17th–18th century), borrowed from Cast. *buscar*, predominates in the rest of the territory.

3.3 Functional and social variation

As is well known, Latin was the language of prestige in all Western Europe during many centuries, and it was widely used in written communication, especially in formal registers. Latin was the basic vehicle for scientific and most administrative communication until not long ago. Although with varying degrees in each language domain, the history of popular vernaculars (that is, all those which were not Latin) has been that of a progressive conquest of spaces of formal expression away from Latin.

Catalan is no exception: it experienced a progressive widening of its domains of usage from the initial Romance words and constructions, attested almost entirely in administrative documents, in the High Middle Ages, to the penetration of Catalan into the most diverse formal registers. This was already observable in the 13th century, and became fully generalized in the ensuing centuries, even while Latin remained the undisputed language of prestige for a long time after. In practice, this penetration into written and formal domains implied, from a sociolinguistic perspective, the continual definition of criteria of linguistic correction, the beginnings of a standardization of Catalan and the development of its function as a social marker; at the same

time, from a linguistic perspective, it implied the creation and spread of expressive resources unique to Catalan.

Due in no small measure to text editions, anthologies and, above all, digital corpora and the use of methods of language history and diachronic sociolinguistics, scholars have begun to acquire an increasingly solid understanding of the diaphasic or functional diversification of Catalan. On this basis, a few classification schemes have been proposed for diverse linguistic registers, as well as their associated text typologies, which have been able to differentiate and classify diverse levels of formality (Miralles 2006; Montoya 2009). A diachronic, global description is yet to be attempted, in which the specific traits of each typology may be described concretely for the lexical domains, building upon the foundation of a matrix of markers of formality.

With specific regard to the lexicon, traces of functional variation can be attested from the first written manifestations in Catalan. These traces amount to early legal expressions, corresponding to a formal register, embedded in administrative texts (feudal oaths, wills, claims) that combine a more or less canonic Latin vocabulary with a diaphasically unmarked Romance vocabulary: *vedar* and *vetar* ‘to impede’, ‘to prohibit’ (11th century), *comonir* ‘to warn, to give notice’ (11th century), *conveniència* ‘agreement’ (11th century).

From the 13th century onwards, there is a steep increase in samples of functional diversification of Catalan, as the typology of texts – both translated from Latin and composed directly in Catalan – becomes more diverse as well (see Miralles’ 2006 typology): we can find a) on the one hand, high-formality texts such as large compilations of legislation (*Usatges* in Barcelona, Bastardas 1984; *Costums* in Tortosa, Massip 1996; the *Furs* of Valencia, Colón/Garcia/Garcia 1980–2007), a vast amount of documents linked to the Royal Chancery (legislative corpus, letters, privileges, etc.) and to the general administration (parliamentary, municipal, feudal, guilds, and ecclesiastical domains), philosophical, historiographical and scientific texts (medicine, veterinary medicine, astronomy, arithmetic), sermons and civil oratory, literature (prose and poetry); and b), on the other hand, there is a geometric progression in less formal testimony, closer to colloquial language: process declarations, private letters, theater and popular literature. Samples of texts from low-formality registers are very easy to find today: everyday conversation or recreated/planned speech (*mediatized oral discourse*) appearing in media, cinema, theater or social networks, are a source of colloquial words such as *acollonir* ‘to lose one’s courage’, *apardalat -ada* ‘dumb’, *bola* ‘lie’, *clapar* ‘to sleep’, *guillar* ‘to leave’, *fotre* ‘to make’, ‘to put’, ‘to damage’, *paio* ‘person, guy, dude’, *xumar* or *xamar* ‘to drink alcohol’ (Salvanyà 2009; Martí Mestre 2011; CCCUB 2008). The vocabulary in this informal register can be diachronically observed in the aforementioned types of text, of which we present a brief sample for illustration (from the Middle Ages to contemporary Catalan):

eixancar ‘to open one’s legs [with a sexual connotation]’ (process, 1277; Ponsoda 1996);
jas ‘here! hold this! [imperative]’, *amiga* ‘lover’, *arrapada* and *arpada* ‘to scratch’,
arpenjar ‘to grab someone violently’, *estiragassar* ‘to pull forcefully’, *aiüixar* ‘to
incite a dog to sth.’ (process, 1358; Miralles 1984);
espatxar ‘to kill’ (process, 1378; Rabella 1998);
acatxar ‘to bend over’ and *rebombori* ‘ruckus’ (satiric poetry, 1490; Gassull 1490);
empapussar ‘to spoon-feed’, ‘to force feed’ (private letter, 1536; Ahumada Batlle
2003);
crepar [*de por*] ‘to die’, *pegar* [*una punyalada*] ‘to stab [to give a stab to someone]’
and *pixarella* ‘fine piss’, ‘light rain’ (journal, 1621; Casas Homs 1975);
femsa ‘manure’ (journal, 1643; Pladevall/Simon 1986);
barrinar ‘to muse, to think’, *guilopo -a* ‘astute’ (novel, 1768; Galiana 21769);
borinot ‘annoying person with no common sense’ (theater, 18th century; Martí
Mestre 1997) and ‘impertinent person’ (theater, 1894; Escalante 1894);
patrocol ‘pile of papers’ (personal journal, 1773; Amat 1987);
xibeca ‘dumb’, *borboll* ‘mess, confusion’, *papu* ‘boogeyman, a figure to frighten
children’, *suca-mulla* ‘bread soaked in wine’ (popular theater, 18th century; Sala-
Valldaura 2007);
esparracat -ada ‘ripped, tattered’, ‘ragged’, *tabola* ‘fuss’ (personal journal, 1825;
Ollé 1981);
bollar ‘to screw up’, *sempentejar* ‘to push’, *mamballetes* ‘applause’, *llemuga* ‘slow,
meticulous person’, *betzol* ‘simple, foolish’, *xarumbada* ‘the act of drinking alco-
hol’ (theater, c. 1865; Penya 1987);
refastinyós -osa ‘fussy’, *tomanyot -a* ‘a gullible person’, *rambolar* ‘to growl’
(costumbrist poetry), 1910; Saisset 1910);
barrina ‘contract’, *fel sofregida* ‘jaundice’, *envergar* ‘to throw’, ‘to place’, ‘to eat’,
xerrera ‘desire to speak’ (traditional tales, 1914; Alcover 1996);
pispar ‘to grab violently’, ‘to steal’ (costumbrist tales, 1919; Ruyra 1919);
budell ‘prostitute’ (detective novel, 1978; Pedroló 1978).

Greco-Latin vocabulary is a fundamental source for the generation of scientific terminology and learned words from the Middle Ages until today; in this field, as has been said before, researchers often invoke the work of Ramon Lull (13th–14th century) and his conscious creation of neologisms for the expression of theological and philosophical concepts necessary to him. This work should not distract from the real engine behind the expansion of the Catalan lexicon in formal registers: the very same expansion of the use of Romance or spoken languages in high-formality registers, that is, the very same fact that these languages, including the Catalan language, in the Middle Ages, progressively became instruments of cognition and communication in ways that went beyond everyday life and colloquial speech. Medieval diglossia, based on a scheme that assigned “Latin to writing and formality” – “Romance to orality and

informality”, started to change. This process is also closely related with the role popular vernaculars play in sociological marking: more formal registers are in truth the re-elaboration of the speech of a concrete and privileged social class and, thus, have a double function as style and social markers (Ferrando/Nicolás 2011, 157).

At the end of the 15th century it is possible to find explicit samples of the speakers’ consciousness of diaphasic variation in Catalan, as well as of the need to regulate it: the highly valued *Regles*, cited above; this may account for the fact that, specifically regarding lexical variation, these norms recommend *latrines* or *privades* for *baça* ‘latrine’, *instrument* for *strument* ‘instrument’, *màrtir* for *martre* ‘martyr’, *consyderació* for *esguard* ‘consideration’; other recommendations in the *Regles* focus on other areas of variation (dialectal and social). This proposal to set a *good* Catalan takes place in a period of a) great elaboration of the formal register under humanist influences, especially the so called *valenciana prosa* (*artitzat* ‘artificial’ model, with a complex syntax which is also influenced by Latin, and a great abundance of learned words Ferrando 1993) and b) debate about the proper usage of spoken languages (*questione della lingua*), in the Catalan domain as well as, above all, in Italy (where it originated), in France or Castile (Nadal/Prats 1996; Badia 1999; Colón/Ferrando 2011).

A very interesting process, from the Lexical-semantic change point of view, is the influence of formal registers in the general language, even more informal registers: traditional words and variants have been substituted with learned alternatives or live on in the periphery or in phraseology: *lleu* ‘lung’ (cited above) for *pulmó*, *darrer -a* ‘last’ for *últim -a*, *lledesme* ‘legitimate’ for *legítim -a*, *soplegar* ‘supplicate’ for *suplicar*, *porgar* ‘purge’ for *purgar*, *regonèixer* ‘recognize’ for *reconèixer*, *esmar* (*aes-*, *as-*) for *estimar* ‘to estimate’. On the other hand, some words of learned origin have entered general usage, frequently with formal or semantic changes: the legal-administrative term *trasllat* (or *trans-*) ‘legal copy of a document’ was adapted according to a predictable Catalan pattern (*tra(n)s-* > *tres-*), to *trellat* or *entrellat*, metonymically to ‘advantage’, ‘sense’ and ‘common sense’ (18th century); *metàfora* ‘metaphor’ has produced, especially in the Balearic Islands and the Valencian region, a series of variants (*mettàfara*, *metàfara*, *mantàfora*, *mantàfola*, *metàcula*) that mean ‘lie, deceit, trick’ (16th century; Martines 2000b, 205); the Latin word of Hellenic origin *phisiognomon* generated the derivate *fisionomia* or, later in the evolution, *fisonomia*, of which *phisomia* appears in the 14th century, and *fesomia* ‘facial features’ in the 15th century, and its usage is generalized in today’s Catalan; church Latin *tu autem* (“Tu autem, Domine, miserere nobis”) evolved into *tuàutem* and *tuacte* ‘main person or thing’, ‘important fact or matter’, ‘plan to carry out something’ (Bastardas 1989).

The process of the generation and maintenance of diaphasic variation we have described, similarly to what happened in adjacent languages, was altered in the Modern Age (and recovered until the 19th–20th century) due to the restrictions Catalan was subjected to in formal domains (with a different intensity in each territory), its limitation to more informal, oral usage, and the interposing of Spanish (Spanish territory, including northern Catalonia until 1659), French (in Northern

Catalonia after 1659), or Italian (in l'Alguer). These languages went on to occupy high formality registers and acted as a source for the generation of learned words, or else they conditioned the development of this process in Catalan. Cultural innovations, often conceptualized in other languages (usually French or English), entered Catalan mediated by Spanish. The latter's prestige rendered it the main vehicle for literary expression, and imposition politics cornered Catalan out of administrative domains (Nueva Planta Decrees, 18th century) and the education system. It was a process aimed to restricting Catalan from formal domains. More on this process in 711 The Growth and Expansion of Catalan (1213–1516).

This sociolinguistic context weakened formal channels of reference, which were used to spread the language model, it colloquialized language usage and intensified dialect diversification. Lexical effects have been described in 2; see also 4. In the latter we tackle the contemporary production of neologisms and terminology.

Some specialists (Mas 2003; Montoya 2009) have properly recognized the noteworthy Catalan tradition that applies Labovian sociolinguistics to dialectal and stylistic variation, and the phenomena that derive from language contact; in contrast, this methodology has been less widely applied to social variation, especially in lexical-semantic domains. The status of Catalan, subjected to a process of minoritization and interference, may have been a condition for this state of affairs: in fact, since the Middle Ages, Spanish (or French or Italian, depending on the territory) has acted as a proper social marker. With varying degrees of intensity in different regions of the Catalan language domains, these languages began to act as instruments of social progress, prestige and social distinction.

Although not widely researched, it seems apparent that Catalan presented diastatic variation in the Middle Ages. Even from a conceptual point of view, between stylistic (or diaphasic) variation, social (or diastratic) variation, and even territorial (or dialectal) variation, there are numerous points of intersection, as concrete social classes do determine the spoken and written formal models. Conversely, colloquial speech and the language spoken by lower, barely educated, classes tend to coincide, as these classes tend not to participate in the construction and use of formal registers. Lastly, different dialectal forms tend to be assigned varying degrees of formality, as well as to different social groups.

It is well established that the *artitzat* style of *valenciana prosa* would have correlated highly with the tastes and expressive mannerisms of the higher classes in the 15th century (Ferrando 1993; Ferrando/Nicolás 2011, 157); likewise, the notions of stylistic, social and dialectal variation may explain the way the *Regles* qualified specific expressions, words or variants as *parlar de baixa sort* 'lower class speech', *parlar de minyons* 'children's speech' or *parlar de dona rustical* 'rustic woman's speech': "*siu-te aquí*" for "*seu-te aquí*" 'sit here'; *de gom a gom* 'completely full'; *tripajoch* 'mess, complication'; *tabustol* 'noise, commotion'.

A good example of vocabulary linked to the domestic domain and affection between noblewomen in the 16th century is found in the letters between a daughter,

Estefania de Requesens, and her mother, Hipòlita de Liori (Ahumada 2003): *estar prenyada* ‘to be pregnant’, *prenyat* ‘pregnancy’ (*mudar de prenyat* ‘to become pregnant again’), *pesar* ‘to regret’, *ànsia* ‘worry’, *descuidar* ‘to neglect, to fail to do something’, *tenir desig* ‘to desire, to look forward to sth.’, *folgar* ‘to cheer up’, *sentir alegria* ‘to feel joy’ (Antolí forthcoming), *voler (gran) bé (a algú)* ‘to wish someone well’, *engruixar* ‘to gain weight (a person)’, *fretura* ‘necessity’ (*passar fretura (d’una cosa)* ‘to suffer the lack of something’), *enyorament* ‘to pine for something’ (*prendre enyorament (d’alguna cosa o d’algú)* ‘to miss’), *trontollat -ada* ‘dizzy’, *desmenjat -ada* ‘lacking appetite’, *atabalat -ada* ‘disconcerted’.

Even despite situational constraints, declarations in lawsuits shed some light onto the vocabulary of various social classes. Miralles (1998) devised an inventory of forms of personal address used by witnesses in 13th and 14th century trials; it includes those employed between persons of the same, low class (*compare* and *companyó* ‘buddy, comrade’, *germà* ‘brother’, *comare* ‘girlfriend, sister’, *fadrí* ‘youth’) and those reserved for persons of a higher social class, or used to express respect (*madona*, *profembra*, *senyora*, for women; *monsenyor*, *monsényer*, *mossèn*, *baró*, for men).

Ancient testimonies display the speakers’ consciousness of a type of lexical variation that is conditioned by the parameter of profession, and even of its value as a social marker and an instrument of the transmission of precise information (specialty language or specific terminology). The guild of *barbers* ‘barbers’ petitioned the *jurats* ‘municipal governors’ of Valencia in 1478 to be equated with the guild of *cirurgians* ‘surgeons’ and, in order to improve their competence in anatomy (*notomia*), they demanded access to cadavers: thus, they could see “*los cosos en les parts dedins com estan organitzats*” (‘the bodies in their internal parts and how they are organized’) and learn the *vocables e noms propis a cascú de aquells* (‘proper names and vocabulary for each of them’) (Rubio 2003, 194).

The evolution of a professional lexicon in Catalan, from the Middle Ages to today, may be studied in abundant documents: advice manuals, guild bylaws and accounting, professional manuals (*menescalia* ‘veterinary medicine’, medicine, stonemasons, agriculture, electronics, gastronomy, etc.), legal processes, diet books, press and advertising, cinema, theater and television and radio shows. Even in marginal social domains we know, for instance, that the *pobla de les fembres pecadrius* ‘town of sinning women’ or the *públic* was the name of the well-known brothel of Valencia, a liberal and cosmopolitan city in the 15th and 16th century.

Documents associated with prostitution in that period reveal that *dones de guany* ‘women of profit’ or *dones que viuen del quest* ‘women who live on soliciting’ were generic terms for prostitutes. *Fembra o dona de cadira* ‘woman of chair’ was the “legal” prostitute, registered with a brothel and paying municipal taxes; in contrast, women who exercised their profession outside of legal regulations were called *fembres de vall* ‘women of the ditch’, *fembres de vila* ‘women of the village, public women’, *dones dissolutes* ‘dissolute women’, *fembres cantoneres* ‘corner women’, *fembres de vora mur* ‘women from the side of the wall’ or *fembres escuseres* ‘secret

women'. Other general terms to express moral or religious repudiation were: *dona deshonest* 'dishonest woman', *dona mundana* 'worldly woman', *dona de mal viure* 'woman of bad living', *fembra pública* 'public women', *dones errades* 'lost or marginalized women', *fembra àvol de son cos* 'a woman who is bad in her body', *fembra pecadriu* 'sinning woman'. Within this domain, the most derogatory terms were *puta* or *bagassa* 'whore', *vil* 'evil', *gossa* or *perra* 'bitch'. The brothel manager was the *hostaler del públic* and the *proxeneta* of a woman was his pimp (Pérez García 1991).

As for the language of different age groups, the following words offer an illustration of their vocabulary: *mam* or *ma* 'water', *ba* or *be* 'kiss', *cocou* or *coco* 'egg', *baba* 'grandmother', *bua* 'damage', *non-non* or *nones* 'sleep', 'slumber', *nyam-nyam* or *mamam* 'to eat', all belong to the language of small children; *penjar* 'to fail (an exam)', *fer campana* or *fer fugina* 'to play hooky', *grillat -ada* 'crazy' may serve to illustrate the language of youth and students.

Parallel to the penetration of formal and specialty words into general usage, expressions associated with specific sociolects and less formal registers have entered general usage. This tendency can be traced all the way to the origins of Romance languages, which emerged from informal registers and popular classes. Let us recall examples already provided (see above), from an informal and familiar (*noi noia* or *al·lot al·lota*), or expressive register (*estripar*, *xafar*), as well as professional domains like agriculture (*trepitjar*, *esguellar* (o *-allar*)); words like these, at various times, have been incorporated progressively into basic Catalan in most or all the territory. Additionally, *be*, probably derived as an onomatopoeia of the cry of a sheep or lamb, has become the name of the animal; *copsar*, originally, 'to catch in flight' has extended into 'to understand'.

4 The Catalan lexicon today and into the future

As we reach the end of this tour through Catalan, we focus on the defining characteristics of the general lexicon of Catalan, paying special attention to the factors that determine its current situation and prospects.

4.1 The standardization of the Catalan lexicon

Catalan transitioned from a state of quick and significant standardization in the Middle Ages, to a state where it required a common normative model for the entirety of its territorial domains and the diversity of uses in a modernizing society, at the turn of the 20th century. In between, there has been a lengthy period of alterations in the social status of the language, as well as interposition and interference of Spanish, French and Italian. The work of Pompeu Fabra, with his *Diccionari general de la llengua catalana* (1932) cradled in the Institut d'Estudis Catalans, became an essential

point of reference in the process of constructing a contemporary, standard Catalan lexicon. The Spanish Civil War (1936–1939) and the long dictatorship of Francisco Franco (1939–1975), as well as the repression of official and public use of Catalan, were a devastating setback in the process to recover a characteristic vocabulary, to spread a reference model and generate original solutions to the new cognitive and communicative needs of that society.

Currently, we are observing a process in which the standard lexical model is being penetrated by significant segments of society: above all, those segments that have received formal education in Catalan. Likewise, there has been an important increase of channels of communication through which speakers can gain access to the standard model: written and audiovisual media, cinema, theater, social media, popular music, even graphics and signage are areas in which, despite the predominance of Spanish (French, or Italian) and English, Catalan presence has been sustained and increasing. Although these improvements have been of varying degrees in the different territories, we can still observe the generalization of words that were previously either only written or belonging to specific geographic varieties: *tardor* ‘autumn’, *xai* ‘lamb’, *amanida* ‘salad’, *mongeta* ‘bean, pea’. Expressions that were previously characteristic of Central and Northern Catalan have advanced and become part of the passive (currently understood but not generally used), sometimes even the active vocabulary as the unified solution in a large part of the territory, for *primavera d’hivern*, *be* or *corder*, *ensalada* and *fesol*. Similarly, words of Castilian origin that were deeply ingrained, have receded significantly: *bocadillo* ‘sandwich’, *cenicero* ‘ashtray’, *grifo* ‘tap’, replaced by *entrepà*, *cendrer* or *aixeta*.

One phenomenon that is worth delving into is the particularist reaction by some territories to this standardization process. This reaction is especially virulent, and it carries specific connotations, in the Valencian region, where lists of alleged “words of Catalan origin” to be avoided have appeared, often contrasted with allegedly genuine Valencian alternatives. Here, these attitudes may serve political aims. Studies show that, even here, the standardization process makes great, uncontroversial strides among formally educated populations (see Segura 2002; 2003; Baldaquí 2002). The foundation of the Acadèmia Valenciana de la Llengua has allowed us to focus on lexical elements which are more markedly linked with regional identity, while avoiding fracturing the unity of the language.

The cultural marketplace and a shared network of several types of media is essential to the maintenance and deepening penetration of the standard model. For any language that aspires to survive in contemporary society, such a model is of the essence as well.

4.2 Lexical-semantic creativity

A further great challenge for Catalan is related to its chances of survival as a differentiated language: lexical-semantic creativity. The secular interference of Spanish (French and Italian), and the current globalization process, are certainly two factors that determine its chances. From this perspective, the important work of centers to produce neologisms and terminology, such as the TERMCAT (722 Terminology and Neology), cannot be stressed enough. Equally important is the role of the education system and mass media. Let us recall that even in the previous century, without the means at the disposal of the speakers, plenty of innovations were transmitted: the old expression *bústia*, today almost forgotten and retained locally as a 'brush (for church alms)', to name a new element, 'letterbox'. The expressions *entrepà*, *barret*, etc. widely replaced the expressions of Spanish origin *bocadillo* 'sandwich' or *sombrero* 'hat'. The current task of spreading neologisms and specialty terminology is fundamental in regenerating the ability of Catalan to generate lexical-semantic innovation. Indeed, as has already been discussed, we need a balance between live, speaker usage and mid-term and long-term language planning processes.

4.3 Conditioned evolution

A further front for the Catalan lexicon, which has been open for centuries (though less so in current times) is the ability to sustain an evolutionary process all on its own. On the one hand, English exerts a worldwide influence both on everyday language and specialized terminology. On the other, the naturalization of Spanish (or French or Italian) among all Catalan speakers, as well as their bilingualization, have reached a previously unheard-of degree. Spanish and Catalan may coexist in one speaker's discourse within the same situation. What is more, due to an important level of exposure to Spanish, it is not unusual for speakers of Catalan to have more (active or passive) knowledge of certain lexical fields in Spanish than in Catalan. Semantic interference from Spanish has not stopped, but rather intensified in younger generations: lexical-semantic changes in Spanish swiftly find themselves reflected in Catalan.

An education system that can compensate for the lacking social presence and communicational competence, as well as the creation of a Catalan cultural sphere, may help to counteract these tendencies.

5 Bibliography

- Adiego, Ignasi-Xavier (2002), *Les llengües preromanes indoeuropees de la Península Ibèrica directament documentades*, in: Joan Rabella (ed.), *Els substrats de la llengua catalana: una visió actual*, Barcelona, Institut d'Estudis Catalans, 33–52.
- Ahumada Batlle, Eulàlia (ed.) (2003), *Epistolaris d'Hipòlita Roís de Liori i d'Estefania de Requesens (segle XVI)*, València, Universitat de València.
- Alcover, Antoni M. (1996), *Aplec de rondalles mallorquines d'En Jordi d'es Racó*, ed. Josep A. Grimalt, Palma de Mallorca, Moll.
- ALDC = Veny, Joan/Pons, Lúcia (2001–2016), *Atles lingüístic del domini català*, 9 vol., Barcelona, Institut d'Estudis Catalans, <http://aldc.espais.iec.cat/> (last accessed: 25.04.2018).
- ALDT = Gimeno Betí, Lluís (1997), *Atles lingüístic de la diòcesi de Tortosa*, Barcelona, Institut d'Estudis Catalans.
- Amat, Rafael d' (1987), *Calaix de sastre*, vol. 1, ed. Ramon Boixareu, Barcelona, Curial.
- Antolí, Jordi M. (2016), *Els verbs de percepció en català antic. Els verbs “veure”, “sentir”, “oir” i “entendre” en els segles XIII–XVI*, Barcelona/Alacant, Publicacions de l'Abadia de Montserrat/Institut Interuniversitari de Filologia Valenciana.
- Antolí, Jordi M. (forthcoming), *Observacions sobre l'expressió dels sentiments en el “Curial e Güelfa”*, in: Anna Maria Babbì/Antoni Ferrando (edd.), *“Curial e Güelfa”. La cavalleria umanística italiana nel XV secolu*, Amsterdam/Philadelphia, Benjamins.
- Badia i Margarit, Antoni (1981), *Gramàtica històrica catalana*, València, 3i4.
- Badia i Margarit, Antoni (1999), *Les “Regles de esquivar vocables” i la “qüestió de la llengua”*, Barcelona, Institut d'Estudis Catalans.
- Baldaquí, Josep M. (2002), *El model lèxic dels estudiants dels programes educatius bilingües*, Catalan Review 16, 29–49.
- Barceló, Carme (1984), *Precisiones sobre algunos arabismos en valenciano y otros nuevos*, Quaderns de Filologia 1, 25–28.
- Barceló, Carme (2011), *Àrab i català: contactes i contrastos*, Barcelona/València, Publicacions de l'Abadia de Montserrat/Institut Interuniversitari de Filologia Valenciana.
- Barri, Montserrat (1999), *Aportació a l'estudi dels gal·licismes del català*, Barcelona, Institut d'Estudis Catalans.
- Bastardas, Joan (ed.) (1984), *Usatges de Barcelona*, Barcelona, Fundació Noguera.
- Bastardas, Joan (1989), *Tu autem*, in: Adolfo Sotelo/Marta C. Carbonell (edd.), *Homenaje al Profesor Antonio Vilanova*, Barcelona, Universitat de Barcelona, 791–804.
- Batlle, Mar, et al. (2016), *Gramàtica històrica de la llengua catalana*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Bosch i Rodoreda, Andreu (2012), *El lèxic alguerès de l'agricultura i la ramaderia entre els segles XVII i XVIII*, Barcelona, Institut d'Estudis Catalans.
- Bruguera, Jordi (1985), *Història del lèxic català*, Barcelona, Enciclopèdia Catalana.
- Cabrè, Teresa M./Domènech, Ona/Estopà, Rosa (edd.) (2014), *Mots nous en català/New words in Catalan*, Amsterdam/Philadelphia, Benjamins.
- Casas Homs, Josep M. (ed.) (1975), *Dietari de Jeroni Pujades*, 3 vol., Barcelona, Fundació Salvador Vives Casajuana.
- CCCUB = Grup d'Estudi de la Variació, *Corpus de Català Contemporani de la Universitat de Barcelona*, Departament de Filologia Catalana, Universitat de Barcelona, 2008, <http://www.ub.edu/ccub/> (25.04.2018).
- CICA = Torruella, Joan, et al. (edd.), *Corpus Textual Informatitzat del Català Antic*, Barcelona, Universitat Autònoma de Barcelona, <http://www.cica.cat/> (25.04.2018).

- CIMTAC = Josep Martines, et al. (edd.), *Corpus Informatitzat Multilingüe de Textos Antics i Contemporanis*, Alacant, Universitat d'Alacant/ISIC-IVITRA.
- Colón, Germán (1976), *El léxico catalán en la Romania*, Madrid, Gredos.
- Colón, Germà (1997), *Estudis de filologia catalana i romànica*, Barcelona/València, Publicacions de l'Abadia de Montserrat/Institut Interuniversitari de Filologia Valenciana.
- Colón, Germà/Ferrando, Antoni (2011), *Les "Regles d'esquivar vocables" a revisió*, València/Barcelona, Institut Interuniversitari de Filologia Valenciana/Publicacions de l'Abadia de Montserrat.
- Colón, Germà/Garcia, Arcadi/Garcia, Vicent (edd.) (1980–2007), *Furs de València*, 11 vol., Barcelona, Barcino.
- Corbera, Jaume (2000), *Caracterització del lèxic alguerès*, Palma, Universitat de les Illes.
- Coromines, Joan (1952), *D'alguns germanismes típics del català*, in: *Mélanges de linguistique et littérature romanes offerts à Mario Roques*, vol. 4, Baden, Éditions Art et Science, 27–37.
- Coromines, Joan (1977), *Mots catalans d'origen aràbic*, in: Joan Coromines, *Entre dos llenguatges*, vol. 3, Barcelona, Curial, 68–177.
- Correa, José A. (2005), *Elementos no indoeuropeos e indoeuropeos en la historia lingüística hispánica*, in: Rafael Cano (ed.), *Historia de la lengua española*, Barcelona, Ariel, 35–57.
- Corriente, Federico (1999a), *Las etimologías árabes de Joan Coromines*, in: Joan Solà (ed.), *L'obra de Joan Coromines, Sabadell*, Fundació Caixa de Sabadell, 67–87.
- Corriente, Federico (1999b), *Diccionario de arabismos y voces afines en iberorromance*, Madrid, Gredos.
- Corriente, Federico (2001), *El romandalusí reflejado por el glosario botánico de Abulxayr*, *Estudios de Dialectología Norteafricana y Andalusí* 2, 91–241.
- Corriente, Federico (2008), *Romania arabica. Tres cuestiones básicas: arabismos, "mozárabe" y "jarchas"*, Madrid, Trotta.
- CTILC = *Corpus Textual Informatitzat de la Llengua Catalana*, Barcelona, Institut d'Estudis Catalans, <http://ctilc.iec.cat> (10.10.2018).
- Cuenca, Maria Josep/Massip, M. Àngels (2005), *Connectors i processos de gramaticalització*, *Caplletra* 38, 259–277.
- D'Aguiló = Aguiló, Marian (1914–1934), *Diccionari Aguiló*, Barcelona, Institut d'Estudis Catalans.
- DCECH = Corominas, Joan/Pascual, José Antonio (1980–1991), *Diccionario crítico etimológico castellano e hispánico*, 6 vol., Madrid, Gredos.
- DCVB = Alcover, Antoni Maria/ Moll, Francesc de Borja (1930–1962), *Diccionari català-valencià-balear*, 10 vol., Palma, Moll.
- De Smet, Hendrik/Verstraete, Jean-Christophe (2006), *Coming to Terms With Subjectivity*, *Cognitive Linguistics* 17, 365–392.
- DECat = Coromines, Joan (*1983–1991), *Diccionari etimològic i complementari de la llengua catalana*, 10 vol., Barcelona, Curial.
- Duarte, Carles/Alsina, Àlex (1984), *Gramàtica històrica del català*, 3 vol., Barcelona,
- Escalante, Eduard (1894), *Colección completa de las obras dramáticas de Don Eduardo Escalante*, 3 vol., València, Domènech.
- Fabra, Pompeu (1932), *Diccionari general de la llengua catalana*, Barcelona, Llibreria Catalònia.
- Ferrando, Antoni (1993), *Sobre una etiqueta historiogràfica de la literatura catalana*, *Caplletra* 15, 11–30.
- Ferrando, Antoni (2011), *L'orientació diatòpica de les "Regles d'esquivar vocables o mots grossers o pagesívols"*, *eHumanista* 18, 316–335, <http://www.ehumanista.ucsb.edu/volumes/18> (last accessed: 10.10.2018).
- Ferrando, Antoni/Nicolás, Miquel (2011), *Història de la llengua*, Barcelona, UOC.

- FEW = Wartburg, Walther von (1922–2002), *Französisches Etymologisches Wörterbuch. Eine Darstellung des galloromanischen Sprachschatzes*, 25 vol., Leipzig, Éditions de Linguistique et de Philologie/Société de Linguistique Romane.
- Fuster, Joan (1976), *La Decadència al País Valencià*, Barcelona, Curial.
- Galiana, Lluís (1769), *Rondalla de rondalles*, València, Monfort.
- Garcia Sebastià, Josep V. (2017), *La gramaticalització de temps ha: de la noció de “temps transcorre-gut” als usos discursius (segles XVI–XX)*, *Zeitschrift für Katalanistik* 30, 77–98.
- Gassull, Jaume (ca. 1490), *Brama dels llauradors de València contra lo venerable mossén Fenollar*, in: Ramon Miquel i Planas (ed.) (1911), *Cançoners satírics valencià dels segles XV y XVI*, Barcelona, Rius.
- Geeraerts, Dirk (1997), *Diachronic Prototype Semantics: A Contribution to Historical Lexicology*, Oxford, Clarendon Press.
- Geeraerts, Dirk (2010), *Theories of Lexical Semantics*, Oxford, Oxford University Press.
- Gilliéron, Jules/Roques, Mario (1912), *Études de géographie linguistique d'après l'Atlas linguistique de la France*, Paris, Champion.
- Gómez, Yorick (2012), *Gli italianismi nel catalano: dizionario storico-etimologico*, Roma, Aracne.
- Gorrochategui, Joaquín (2002), *Las lenguas de los Pirineos en la antigüedad*, in: Joan Rabella (ed.), *Els substrats de la llengua catalana: una visió actual*, Barcelona, Institut d'Estudis Catalans, 75–102.
- Grice, Paul H. (1975), *Logic and Conversation*, in: Deborah Schiffrin (ed.), *Studies in the Way of Words*, Harvard, Harvard University Press, 22–40.
- Horn, Laurence R. (1984), *Toward a New Taxonomy For Pragmatic Inference: Q-Based and R-Based Implicature*, in: Deborah Schiffrin (ed.), *Meaning, Form, and Use in Context: Linguistic Applications*, Washington, Georgetown University Press, 11–42.
- Jaime, Joan M. (2015), *El lèxic d'origen germànic en el llatí medieval de Catalunya*, Barcelona, Universitat de Barcelona.
- Jaime, Joan M. (2017), *Germanismos del catalán a partir del latín medieval*, *Revista de filología y lingüística de la Universidad de Costa Rica* 43, 103–114.
- Lamuela, Xavier/Murgades, Josep (1984), *Teoria de la llengua literària segons Fabra*, Barcelona, Quaderns Crema.
- Levinson, Stephen C. (1995), *Three Levels of Meaning*, in: Frank R. Palmer (ed.), *Grammar and Meaning: Essays in Honour of Sir John Lyons*, Cambridge, Cambridge University Press, 90–115.
- LEXDIALGRAM = Perea, Maria-Pilar (coord.), *Lexdialgram. Portal de lèxics i gramàtiques dialectals del català del segle XIX*, Barcelona, Universitat de Barcelona, <http://www.ub.edu/lexdialgram/> (last accessed: 10.10.2018).
- Martí Mestre, Joaquim (1997), *Literatura de canya i cordell al País Valencià: els col·loquis de temàtica jocosa i satírica: edició i estudi lingüístic*, València, Denes.
- Martí Mestre, Joaquim (2011), *Diccionari històric del valencià col·loquial*, València, Universitat de València.
- Martines, Josep (1997), *El sufix -aire al País Valencià*, in: *Miscel·lània Germà Colon*, vol. 7, Barcelona, Publicacions de l'Abadia de Montserrat, 229–262.
- Martines, Josep (2000a), *El canvi lèxic en català (ss. XVI–XX). Una aproximació des de la lexicologia diacrònica cognitiva (I). Les novetats i la llengua catalana*, in: Lluís B. Polanco (ed.), *Jornades de la Secció Filològica de l'Institut d'Estudis Catalans a Elx i a la Universitat d'Alacant: 16 i 17 d'octubre de 1998*, Barcelona, Institut d'Estudis Catalans, 35–65.
- Martines, Josep (2000b), *El valencià del segle XIX. Estudi lingüístic del “Diccionario Valenciano” de Josep Pla i Costa*, Barcelona/Alacant, Publicacions de l'Abadia de Montserrat/Institut Interuniversitari de Filologia Valenciana.
- Martines, Josep (2002), *L'aragonès i el lèxic valencià. Una aproximació*, *Caplletra* 32, 157–201.

- Martines, Josep (2009), *El contacte del català amb la llengua dels aragonesos al segle XIII al País Valencià: influència sobre el lèxic*, Caplletra 46, 61–88.
- Martines, Josep (2011), “Punxar” i família, *mossarabismes del català?*, Caplletra 51, 204–235.
- Martines, Josep (2012a), *Aproximació a les novetats lèxiques i semàntiques del “Curial e Güelfa”*, in: Antoni Ferrando (ed.), *Estudis lingüístics i culturals sobre “Curial e Güelfa”, novel·la cavalleresca anònima del segle XV en llengua catalana*, Amsterdam/Philadelphia, Benjamins, 941–998.
- Martines, Josep (2012b), *Història del lèxic i contacte de llengües. El català i l'aragonès al País Valencià a l'edat mitjana: un tast lèxic*, in: Gloria Clavería et al. (edd.), *Historia del léxico: perspectivas de investigación*, Madrid/Frankfurt am Main, Iberoamericana/Vervuert, 127–166.
- Martines, Josep (2012c), *El valencià del segle XIX. El lèxic: l'aportació del “Diccionario valenciano” de Josep Pla i Costa*, Barcelona, Institut d'Estudis Catalans.
- Martines, Josep (2013), *El verb “estimar” i l'amor hereós i Joan Roís de Corella: un acostament segons la pragmàtica diacrònica*, Afers 76, 717–740.
- Martines, Josep (2015), *Diacronia i neologia: canvi semàntic, subjectivació i representació del pensament. El català “esmar”, des de ‘taxar’ fins a ‘inferir’ i ‘imaginar’ i més enllà*, Caplletra 59, 221–248.
- Martines, Josep (2018), *La història del lèxic i els corpus textuais i lexicogràfics: una ullada sobre “escombrar” i “agranar”*, in: Maria-Pilar Perea/Àngels Massip (edd.), *Noves aproximacions a la lexicografia dialectal*, Barcelona, Universitat de Barcelona, 87–127.
- Martines (forthcoming), *Precisions sobre els elements lingüístics italians del “Curial e Güelfa”: lèxic i fraseologia*, in: Anna Maria Babbi/Antoni Ferrando (edd.), *“Curial e Güelfa”. La cavalleria umanística italiana nel XV secolo*, Amsterdam/Philadelphia, Benjamins.
- Martines, Josep/Montserrat, Sandra (2014), *Subjectivació i inferència en l'evolució semàntica i en l'inici de la gramaticalització de “jaquir” (segles XI–XII)*, Caplletra 56, 185–211.
- Martínez, Caterina (2017), *Evolució i procés de gramaticalització del marcadore discursiu “nores-menys” en català antic*, Zeitschrift für Katalanistik 30, 53–76.
- Martos, Josep Lluís (ed.) (2001), *Les proses mitològiques de Joan Roís de Corella*, Barcelona/Alacant, Publicacions de l'Abadia de Montserrat/Institut Interuniversitari de Filologia Valenciana.
- Mas, Antoni (2003), *La sociolingüística històrica: una alternativa a l'anàlisi del canvi lingüístic*, Noves SL. Revista de sociolingüística 3, http://www6.gencat.net/llengcat/noves/hm03tardor/mas1_2.htm (last accessed: 10.10.2018).
- Massip, Jesús (ed.) (1996), *Costums de Tortosa*, Barcelona, Fundació Noguera.
- Milá y Fontanals, Manuel (1861), *De los trovadores en España: estudio de lengua y poesía provenzal*, Barcelona, Librería de Joaquín Verdager.
- Miralles, Joan (1979), *Sobre l'ús lingüístic en les viles medievals mallorquines. Els llibres de Cort Reial*, in: *Actes del Cinquè Col·loqui de l'Associació Internacional de Llengua i Literatura Catalanes*, Barcelona, Publicacions de l'Abadia de Montserrat, 511–534.
- Miralles, Joan (1984), *Un llibre de cort reial mallorquí del segle XIV (1357–60)*, 2 vol., Palma de Mallorca, Moll.
- Miralles, Joan (1998), *Per a una tipologia del català col·loquial a l'Edat Mitjana*, in: *Actes de l'Onzè Col·loqui Internacional de Llengua i Literatura Catalanes*, vol. 1, Barcelona, Publicacions de l'Abadia de Montserrat, 227–261.
- Miralles, Joan (2006), *Antologia de textos de les Illes Balears*, 6 vol., Barcelona/Palma Publicacions de l'Abadia de Montserrat/Institut d'Estudis Balearics.
- Moll, Francesc de Borja (1952), *Gramàtica històrica catalana*, Madrid, Gredos.
- Montoya, Brauli (2009), *Tipologia textual i de registres en el català antic*, in: Manuel Pérez Saldanya/Josep Martines (edd.), *Per a una gramàtica del català antic*, Alacant, Institut Interuniversitari de Filologia Valenciana, 73–75.

- Montserrat, Sandra (2007), *La semàntica diacrònica cognitiva: una aplicació a propòsit de “venir”, “arribar” i “aplegar” (segles XII–XVI)*, Barcelona/Alacant, Publicacions de l'Abadia de Montserrat/Institut Interuniversitari de Filologia Valenciana.
- Nadal, Josep/Prats, Modest (1982–1996), *Història de la llengua catalana*, 2 vol., Barcelona, Edicions 62.
- Ollé, Josep M. (1981), *Successos de Barcelona (1822–1835)*, Barcelona, Curial.
- PALDC = Veny, Joan (2007–2017), *Petit atlas lingüístic del domini català*, Barcelona, Institut d'Estudis Catalans.
- Pedrolo, Manuel de (1978), *Aquesta nit tanquem*, Barcelona, Edicions 62.
- Penya, Pere d'Alcàntara (1987), *Teatre*, ed. Gabriel Janer Manila, Palma de Mallorca, Biblioteca Bàsica de Mallorca.
- Pérez García, Pablo (1991), *Un aspecto de la delincuencia común en la Valencia pre-agermanada: la “prostitución clandestina” (1479–1518)*, Revista de historia moderna 10, 11–41.
- Pérez Saldanya, Manuel (2015) *Les construccions casuais en català: classes i nexes que les introdueixen*, Els Marges 105, 10–38.
- Pitarch, Vicent (2001), *Llengua i església durant el barroc valencià*, Barcelona/València, Publicacions de l'Abadia de Montserrat/Institut Interuniversitari de Filologia Valenciana.
- Pladevall, Antoni/Simon, Antoni (1986), *Guerra i vida pagesa a la Catalunya del segle XVII*, Barcelona, Curial.
- Pons, Jaume (2018), *Canvi semàntic i gramaticalització dels verbs “metre” i “posar”: un acostament segons la semàntica cognitiva diacrònica*, Alacant, Universitat d'Alacant.
- Ponsoda, Joan J. (1996), *El català i l'aragonés en els inicis del Regne de València segons el “Llibre de Cort de Justícia” de Cocentaina (1269–1295)*, Alcoi, Marfil.
- Rabella, Joan Anton (1998), *Un matrimoni desavingut i un gat metzinat. Procés criminal barceloní del s. XIV*, Barcelona, Institut d'Estudis Catalans.
- Rafel, Joaquim (1996), *El “Diccionari de l'Institut” i el “Diccionari Fabra”*, in: *Estudis de lingüística i filologia oferts a Antoni M. Badia i Margarit*, vol. 3, Barcelona, Publicacions de l'Abadia de Montserrat, 217–269.
- Ramos, Rafael (2016), *El cambio semántico del verbo “pegar” en catalán (siglos XIII–XXI)*, Zeitschrift für romanische Philologie 132, 669–692.
- Rubio Vela, Agustín (2003), *Epistolari de la València Medieval*, vol. 1, València/Barcelona, Institut Interuniversitari de Filologia Valenciana/Publicacions de l'Abadia de Montserrat.
- Ruyra, Joaquim (1919), *La parada*, Barcelona, Editorial Catalana.
- Saisset, Albert (1910), *Perpinyanenques*, Barcelona, L'Avenç.
- Sala-Valldaura, Josep M. (ed.) (2007), *Teatre burlesc català del segle XVIII*, Barcelona, Barcino.
- Salvanyà, Jaume (2009), *Diccionari del català col·loquial*, Barcelona, Enciclopèdia Catalana.
- Sánchez López, Elena (2015), *Phraseologization as a process of semantic change*, Catalan Journal of Linguistics 14, 59–177.
- Sánchez Moreno, Eduardo (ed.) (2007), *Protohistoria y antigüedad de la Península Ibérica*, 2 vol., Madrid, Sílex.
- Sanchis, Manuel (1980), *Aproximació a la història de la llengua catalana*, Barcelona, Salvat.
- Segura, Carles (2002), *Diversitat dialectal i estandarització en el valencià meridional: èxit o fracàs dels models lingüístics a l'escola*, in: Emili Casanova (ed.), *Estudis del valencià d'ara: actes del IV Congrés de Filologia Valenciana del 20 al 22 de maig de 2000: en homenatge al Doctor Joan Veny*, València, Denes, 581–594.
- Segura, Carles (2003), *Variació dialectal i estandardització al Baix Vinalopó*, Barcelona/València, Publicacions de l'Abadia de Montserrat/Institut Interuniversitari de Filologia Valenciana.
- Sentí, Andreu (2017), *Modalitat i evidencialitat en català antic. Un acostament cognitiu a les perífrasis verbals amb “deure” i amb “haver”*, Barcelona/València, Publicacions de l'Abadia de Montserrat/Institut Interuniversitari de Filologia Valenciana.

- TLV = Guardiola, M. Isabel (2004), *Tresor lexicogràfic valencià (1543–1880)*, Alacant, Universitat d'Alacant.
- Traugott, Elizabeth (1989), *On the Rise of Epistemic Meanings in English: An Example of Subjectification in Semantics*, *Language* 65, 31–55.
- Traugott, Elizabeth (2012), *Pragmatics and Language Change*, in: Keith Allan/Kasia Jaszczolt (edd.), *The Cambridge Handbook of Pragmatics*, Cambridge, Cambridge University Press, 549–566.
- Traugott, Elizabeth/Dasher, Richard (2002), *Regularity in Semantic Change*, Cambridge, Cambridge University Press.
- Velaza, Javier (2002), *Darrers avenços en la investigació sobre la llengua ibèrica*, in: Joan Rabella (ed.), *Els substrats de la llengua catalana: una visió actual*, Barcelona, Institut d'Estudis Catalans, 11–32.
- Veny, Joan (*1983), *Els parlars catalans*, Palma de Mallorca, Moll.
- Veny, Joan (2002), *Sobre el valencià “gemecar”, ‘gemegar’*, *Caplletra* 32, 143–155.
- Veny, Joan (2007), *El “Diccionari General de la Llengua Catalana”: precedents, posterioritat, dialectalismes*, in: Joan Solà/Jordi Mir (edd.), Pompeu Fabra, *Obres completes*, vol. 5, *Diccionari general de la llengua catalana*, Barcelona/València/Palma, Proa/Edicions 3i4/Moll, 41–76.
- Veny, Joan (2011), *Sobre el mossarabisme alatxa (Sardinella aurita)*, *Caplletra* 51, 185–203.
- Veny, Joan (2016), *De geolingüística i etimologia romàniques*, Barcelona, Universitat de Barcelona.
- Veny, Joan/Massanell, Mar (2015), *Dialectologia catalana. Aproximació pràctica als parlars catalans*, Barcelona, Universitat de Barcelona/Universitat d'Alacant/Universitat de València.
- Villar, Francisco (2002), *Indoeuropeos y no indoeuropeos en Cataluña y el Noreste hispano*, in: Joan Rabella (ed.), *Els substrats de la llengua catalana: una visió actual*, Barcelona, Institut d'Estudis Catalans, 53–74.
- Villar, Francisco (2008), *Joan Coromines y los substratos prerromanos de la Península Ibérica*, in: Antoni M. Badia/Joan Solà (edd.), *Joan Coromines, vida y obra*, Madrid, Gredos, 368–399.
- Wierzbicka, Anna (2006), *English: Meaning and Culture*, Oxford, Oxford University Press.

Jens Lüdtke

7.2 Word-Formation

Abstract: The aim of the present outline is to show how a semantic approach can provide an adequate account of the formation of complex words in Catalan. Since form serves to express content and has no independent function, it seems quite natural to base word-formation on semantic considerations. There is no one-to-one correspondence between form and meaning; for example, a suffix may or may not change the word class. Thus, it is not possible to present word-formation coherently on both a morphological and a semantic basis at the same time. Similar meanings can be expressed by prefixes and suffixes, intensification by *re-* or *-íssim*, for instance, in *sec* 'dry' → *ressec* 'very dry' and *bo* 'good' → *boníssim* 'very good', and suffixes can be used to convey meanings that are as different as those expressed by *-dor* in *treballar* 'to work' → *treballador* '(hard-)working' as well as 'worker' and in *dormir* 'to sleep' → *dormitori* 'bedroom'. This overview of Catalan word-formation aims to show that a content-based analysis can yield significant results.

Keywords: semantics, morphology, transposition, modification, compounding

1 Some basic approaches to word-formation

Word-formation is concerned with the patterns used by speakers to form new words which are analysable both formally and semantically (Marchand ²1969, 2). At the same time, we should bear in mind that there are certain prescriptive issues relevant to Catalan word-formation (Rull 2004), such as pressure from Spanish word-formation patterns, in order to make sure that only genuine Catalan words are taken into account. Patterns involving a base and an affix, which may be a prefix or a suffix, are described as derivation, a base which changes its word class is an example of conversion, while joining one base to another is known as compounding.

In awareness of the fact that current research into Catalan is morphologically biased, as it is synthesized by the Institut d'Estudis Catalans' recent *Gramàtica de la llengua catalana* – where word-formation is treated as 'lexical morphology' (GIEC 2016, 137–146), which, peculiarly, does not exclude the term 'word-formation' but does not treat it as a separate field –, the present overview attempts to classify Catalan word-formation patterns on a semantic basis, since there is nothing more fundamental in language than meaning and any language should be described from multiple approaches. This topic has been discussed in the heritage of the external argument structure in Gràcia (1995), but not with respect to the internal grammar.

<https://doi.org/10.1515/9783110450408-013>

tical or semantic structure, and it has not been generalised to apply to the whole field of Catalan word-formation.

A morphological approach tends to emphasise the combination of forms and stresses formal differences which are not made by the language: the same suffix, e.g. the diminutive suffix *-et* may apply to a nominal, adjectival, adverbial and verbal pattern which should not be separated, because its basic meaning is the same. A morphological description is more fragmented than a semantic one and is based on classification criteria which do not simultaneously take semantic issues into account, for instance in the case of prefixation, where the distinction between stressed and unstressed prefixes means that unstressed prefixes fall into the category of prefixation and stressed prefixes into that of compounding (Bernhard s.a.; Blasco Ferrer 1984; Mascaró 1985, 31–34; Cabré 2002, 734–739; Gràcia 2002, 785–796).

The lexical meaning of Catalan compounds and derivatives is often taken for granted in publications addressed by native authors to native speakers, but a non-native cannot rely on an implicit linguistic competence and has to explicitly consider the semantic structure if he or she hopes to explain the use of complex words correctly. Based on the principle that complex words must be analysable synchronically in present-day Catalan, the product of a process must be broken down into form and meaning, and this relationship must be made explicit in the form of a paraphrase that gives an account of either the pattern meaning or a lexicalised meaning, i.e. a more general meaning which describes the pattern meaning and a specific meaning which corresponds more or less to a lexicographical definition. This approach is confirmed authoritatively by Fabra in his dictionary (*DGLC*, 1932) and is acknowledged by the current lexicography of the IEC (*DIEC*). Fabra's paraphrases, which will be reduced to their essential elements and cannot be designed to be idiomatic, lend strong support to the present semantic analysis. Linguists should not be afraid to render a simple pattern meaning through a simple paraphrase. However, this procedure is rarely applied to word-formation with a Greek or Latin basis since this type of word-formation is more concerned with the question of analysability than with semantics and will therefore be treated along with the Catalan types in this overview (see Rull 2004; *GIEC* 2016, 463–467) whenever it is relatively transparent. The present approach is expounded in greater detail in Lüdtke's book on Romance word-formation (2011), which also presents the basic semantic issues in the theory of the field.

Word-formation is a particular form of grammaticalisation that appears in three basic types of process and implies the creation of a word based on one or two elements, i.e. transposition, modification and compounding. In the first case, we distinguish between the implication of a grammatical function at the sentence level, or "actual" function, in transposition, i.e. *per què* 'why' → *el perquè* 'the why, the reason', *dubtar* 'to doubt' → *dubte* 'doubt', *fort*, *-a* 'strong' → *fortament* 'very (much)', *de València* → *valencià* (2). In the case of modification, which includes an "inactual" function, grammaticalisation operates on an element at the sub-sentence level, i.e.

the formation of diminutives in *casa* ‘house’ → *caseta* ‘little house’ or *blanc* ‘white’ → *blanquinós* ‘whitish’, the formation of collective nouns in *branca* ‘branch’ → *brancatge* ‘branches’, and repetition in *venir* ‘to come’ → *revenir* ‘to come back’ (3). In the third case, a lexeme is combined with two elements, either with a generic element in *fusta* ‘wood (substance)’ + ‘somebody’ expressed by *-er* → *fuster* ‘joiner’, *menjar* ‘to eat’ + ‘something’ represented by *-dor* → *menjador* ‘dining-room’ (4.1), or two lexemes are joined as in *paper moneda* ‘paper money’ and *catalanoparlant* ‘Catalan-speaking’ to form a compound (4.2). “Lexeme” is used here for that part of a word that has lexical meaning, for example *dubt-*, *fust-* or *-venir* in *dubte*, *fuster* and *revenir*.

In comparison, the morphological or material approach may be summarised as conversion or zero-derivation (*el perquè*), suffixation (*caseta*, *fuster*, *menjador*, *blanquinós*), prefixation (*revenir*, *ressec* ‘very dry’) and compounding (*paper moneda*, *catalanoparlant*). The most striking difference between a morphological and the present semantic approach appears in the concept of compounding, which here is based on meaning and form and is not restricted to form only (Staib 1988). It goes without saying that both approaches admit variations which cannot be discussed here.

In principle, the present approach implies the separation of grammatical processes that may be lexicalised, e.g. the past participle in *passat* ‘past’ and the noun phrase *ull de bou* ‘porthole’, from genuine word-formation and attempts to look past the products of word-formation by starting from looking at the processes behind them (cf. 4.2).

The morphologically oriented reader should have little trouble finding the types he or she is looking for. Since the complex morphology of the Catalan word is dealt with in 15.1 Word Classes, Inflectional Categories and Paradigms, it will be assumed here. English equivalents will not be given for every word, even if the meaning of materially similar words may be slightly different and the cognate forms are therefore misleading.

The references in brackets are sources as well as suggestions for further reading and do not necessarily support the interpretation given in the text (Fabra 1956, 108–159; Badia Margarit 1962, vol. 2, 287–398; Mascaró 1985, 9–84; Cabré/Rigau 1986, 49–156; Cabré 1994; López del Castillo 2002; Solà et al. 2002; Bruguera i Talleda 2006; GIEC 2016, 137–146, 389–467).

2 Transposition

Transposition, my term for Coseriu’s “development” (1978, 250, 255–256), is based on meanings which are analogous to grammatical functions such as predicate, subject, object and other complements, and produces words by means of conversion and suffixation. Patterns include deverbal, deadjectival and denominal predicate nouns, agent nouns and place nouns; deverbal and denominal adjectives, etc. Also falling

under this type of word-formation is conversion, which produces words by a mere change of word class.

2.1 Conversion

Conversion is sometimes confused with ellipsis, e.g. in *abric impermeable* → *impermeable* ‘raincoat’, but it is a process of its own. The simplest type is the change of word class without any inflection in the case of invariable word classes, as in *per què* ‘why’ → *el perquè* ‘the reason’, *el jo* ‘the self; ego’, *el sobre* ‘top; envelope’, *el sota* ‘the dive’, *el dins* ‘the inside’, *el sí* ‘the yes’. Conversion by selection of a morpheme for the base can apply to the infinitive morpheme as in *esmorzar* ‘to breakfast’ → *l’esmorzar* ‘the breakfast’, *ésser* ‘to be’ → *l’ésser* ‘the existence’, the present participle as in *contribuir* ‘to contribute’ → *contribuent* ‘one who contributes, taxpayer’, *estimulant* ‘stimulant’ or the past participle as in *fer* ‘to do’ → *el fet* ‘the act, the fact’, *escriure* ‘to write’ → *l’escrit* ‘the writing, the document’. What I call conversion by change of paradigm is controversial, since this type is also considered as zero-derivation and is documented in *dubtar* → *el dubte*, *marxar* ‘to march’ → *la marxa* ‘the march’. However, both the verb and the noun are sufficiently marked by different sets of morphemes and there is no need to assume the existence of a zero-morpheme. Curiously, the “theoretical dubiousness of speaking of zero affixes” in English (Bauer 1983, 33) is enhanced by the fact that morphology, both grammatical and derivational, was reduced in the transition from Old to Modern English. The approach advocated here is the use of paraphrases which may represent the linguistic consciousness of speakers. Besides, it is not necessary for us to settle descriptive problems here whose status is ambiguous or fuzzy in a particular language (cf. Mascaró 1985, 37–39; GIEC 2016, 410–412; and the discussion in Rull 2016).

2.2 Predicate nouns, or predicative nominalisations

The nouns of this type are currently paraphrased by ‘the action, act or fact of V-ing’ (V = verb), though ‘V-ing’ would be sufficient. The main features of this type include the fact that the process always results in a noun and that it is based on a predicate. The nouns ‘action’, ‘act’ and ‘fact’ or others only imply a subclassification of the underlying verbs, adjectives and nouns. 2.2.3 and 2.2.4 will show that the types *homenia* and *bastonada* are considered transpositions due to the fact that they transpose a predicate + noun (Lüdtke 1978, 240–242).

2.2.1 Deverbal predicate nouns

This type is formed by conversion, as in *assajar* → *assaig* ‘essay, practice’, *desitjar* → *desig* ‘wish’, *escanejar* → *escaneig* ‘scanning’, *perdonar* → *perdó* ‘pardon’, *remolcar* → *remolc* ‘drag’, *treball* ‘work’, *trepitjar* → *trepig* ‘tread’, *contestar* ‘reply’, *neteja* ‘clean (s)ing’, by the feminine form of the past participle in *baixada* ‘descent’, *caure* → *caiguda* ‘(sudden) fall’, *durada* ‘duration’, *pensada* ‘idea’, or by adding suffixes such as *-ment*, *-ció*, *-ança/-ença*, *-atge*, *-dura*, *-dissa* and some less frequent suffixes in *acabament* ‘end’, *enterrament* ‘burial, interment’, *engrandiment* ‘enlargement’; *contestació* ‘answer’, *equivocació* ‘error, mistake’, *indicació* ‘announcement, instruction, reference’, *reparació* ‘repair, overhaul’; *ensenyança* ‘instruction, training, teaching’; *buidatge* ‘emptying’, *embalatge* ‘packing’, *reciclatge* ‘recycling’, *sondatge* ‘sounding’; *dauradura* ‘gilding’; *aixafadissa* ‘crushing, squashing’, *clamadissa* ‘lamentation, wailing’, *cridadissa* ‘shouting, yelling’, *trencadissa* ‘breaking’; *escampall* ‘scattering’; *mengera* ‘appetite’, *pixera* ‘pee’, *xerrera* ‘pleasure to chatter’; *ruixar* → *ruixim* ‘drizzle’; *amor* ‘love’, *cremor* ‘burning, heat’ (Lüdtke 1978, 220–240; GIEC 2016, 398–400).

2.2.2 Deadjectival predicate nouns

This type corresponds to the paraphrase *ser cert* ‘to be certain’ → *certesa* ‘the state of being certain’ or, preferably, ‘the being certain’. Other nouns are related to the semantics of predicative adjectives, e.g. in *feblesa* ‘the (quality or condition of) being weak’, *abstractesa*, *bellesa* ‘beauty’. Other suffixes include *-dat/-tat*, *-ància/-ència*, *-ia*, *-ícia*, *-itud*, *-or*, *-eria*, *-ària*, *-isme*: *claredat* ‘clearness’, *netedat* ‘clean(li)ness’, *profunditat* ‘depth’, *vanitat*; *ignorància*; *negligència*; *audàcia* ‘boldness’; *magrícia* ‘meagreness’; *exactitud* ‘exactness’; *calentor* ‘warmth’, *roig*, *roja* → *rojoir* ‘redness’; *boig*, *boja* → *bogeria* ‘madness’; *alegria* ‘joy, gladness’, *cortès*, *-esa* → *cortesia* ‘politeness’; *gran* → *grandària* ‘size, largeness, spaciousness’, *eròtic* → *erotisme* (Lüdtke 1978, 242–253; GIEC 2016, 398).

2.2.3 Denominal predicate nouns

Examples of paraphrases are *home* → *homenia* ‘the (state of) being a man, in general or as opposed to childhood or to womanhood’ and *poncella* → *poncellatge* ‘girlhood, maidenhood’, that is to say, for instance, ‘the state or time of being a girl’ or simply ‘the being a man, a girl, etc.’. Accordingly, the most frequent meanings which appear in denominal predicative nouns, besides the predicative meaning, are of place and time. The suffixes *-at*, *-atge*, *-ia* and *-eria* belong to this class: *bisbat* ‘bishopric, diocese’, *comtat* ‘county’; *vassallatge* ‘the being a vassal’; *alcaldia* ‘post of mayor’; *beneiteria* ‘stupidity, a foolish thing’ (Lüdtke 1978, 253–257; Cabré 2002, 747–749).

2.2.4 The types *bastonada* and *criaturada*

These types denote an action which is not expressed by an overt verb. The first example refers to blows, stabs, hits and the like with or on something as in *bastó* ‘stick’ → *bastonada* ‘blow with a stick’ or *clatell* → *clatellada* ‘blow behind the neck’, and to a typical action or behaviour of a person, i.e. *criatura* ‘child, tomfool’ → *criaturada* ‘the action or behaviour of a tomfool, tomfoolery, childishness’ (GIEC 2016, 401).

2.3 Verbalisation

Verbalisation is a form of transposition based on nouns (Grossmann 1989), adjectives and adjuncts (Cabr  2002, 751–756). The underlying nouns are converted into verbs in the function of object as in *flor* ‘flower’ → *florir* ‘to flower’ or of adjunct as in *esper * ‘spur (as an instrument)’ → *esperonar* ‘to spur’, or of subject by means of a suffix as in *capitanejar* ‘to be a captain’ (GIEC 2016, 404–406).

2.3.1 Subject-oriented denominal verbalisation

A possible paraphrase is ‘to be an assassin, to act as an ~’ → *assassinar*, which refers to a subject or the predicative use of an animate noun, as in *pilot* → *pilotar*, *senyorejar* ‘to be/act as/like a lord’. Nouns that refer to animals frequently designate typical behaviour of that animal, for example *papall *, *-ona* → *papallonejar* ‘to move like a butterfly’.

2.3.2 Object-oriented denominal verbalisation

This type forms verbs by conversion (*arrel* → *arrelar* ‘to (cause to) take root’, *creu* → *creuar* ‘to cross’, *esper * → *esperonar*, *vedell*, *-a* → *vedellar* ‘to bring forth a calf’, *doctorar* ‘to confer the degree of doctor on sb’, *angoixa* → *angoixar* ‘to cause anguish’) or suffixation (*aire* → *airejar* ‘to aerate’, *feina* → *feinejar* ‘to be busy with sth’). The object is given in the following paraphrases: *an lisi* → *analitzar* ‘to make an analysis of sth’, *tros*, *-ossos* → *trossejar* ‘to make pieces, i.e. to divide into pieces’, *xiuxiuejar* ‘to make *xiu-xiu*, i.e. to whisper’.

2.3.3 Adjunct-oriented verbalisation

Some verbs of this type are derived directly from a noun with or without a suffix (*escombrar* ‘to clean with a broom’, *finestrear* ‘to be at the window’, *costejar* ‘to sail by the coast’).

The more frequent and highly controversial type is based on preposition + noun. This was called parasyntesis by Darmesteter (1875, 82–83), who assigns the prefixes which appear in these forms to the classes of adverbs or prepositions (1875, 95). This was also the view taken by Fabra, as seen in his paraphrases: *embossar* ‘posar *a la bossa*, to put *into* one’s *pocket*’, and the type preposition + noun is rightly considered by Mascaró (1985, 32) as a case of conversion. This interpretation is less evident in Catalan than in other Romance languages, since the prefix *en-/em-* corresponds to both the prepositions *a* and *en*. It is even less so when the linearity of the base and the result of the process are inverted as in *desgranar/esgranar* ‘to extract *the grain/seed* (A) *from* sth (B)’ (‘se separa *la grana* (del cotó) (A) *de* les fibres (B)’). The correspondence between the prefixes *des-* and *es-* and the preposition *de* is rather opaque, but still functions. This relationship can only be explained by progress in the analysis of prepositions which establishes the relational function between two elements (Lang 1991; Weidenbusch 1993; cf. GIEC 2016, 437), not only the one that follows, ‘of the fibres (B)’, but also the one that precedes, ‘the grain (A)’ in the example.

The adjuncts may denote an instrument, a place, a time or a manner in *abordar* ‘to board’, derived from *a bord* ‘on board’, *embotellar* ‘to bottle’, *embutxacar* ‘to pocket’, *endinsar* ‘to put into’, *engabiar* ‘to cage (up)’, *ensabonar* ‘to soap’, *esteranyinar* ‘netejar de teranyines, to clear of cobwebs’, *descoratjar* ‘to discourage’, *desfullar* ‘llevar les fulles (d’una planta, branca, etc.)’. Grossmann (1994) accounts for the “ingressive” meaning in *engabiar* and the “egressive” meaning in *desfullar*, an approach that would make more sense through the interpretation of *en-*, *es-*, etc. as “prepositional prefixes” (Lang 1991). A detailed analysis will have to account for various degrees of motivation from transparency (*embotellar*) to relative opaqueness (*esporuguir* ‘to intimidate’).

2.3.4 Subject-oriented deadjectival verbalisation

This type is given by conversion in *baixar* ‘to go down’ and by suffixation in *agregar* ‘to be sour, to have a sour taste’, *fosquejar* ‘to grow dark’.

2.3.5 Object-oriented deadjectival verbalisation

This type is expressed by conversion in *alegrar* ‘to rejoice’, *espessir* ‘to thicken’, *estretir* ‘to tighten’ and by suffixation in *justificar* or *simplificar* (GIEC 2016, 406).

2.3.6 The adjunct-oriented deadjectival verbalisation

This type is based on preposition + adjective and belongs in the same category as *embossar* and *esgranar* (2.3.3): *abaixar* ‘fer descendir a baix’, *afeblir* ‘canviar a feble, to enfeeble’, *allargar* ‘to lengthen’, *assegurar* ‘to secure’, *embrutar* ‘to soil, to dirty’, *engrandir* ‘to enlarge’, *enterbolir* ‘to make turbid’, *entristir* ‘to sadden’, *esporuguir* (GIEC 2016, 425, 437, 438).

2.4 Adjectivalisation

2.4.1 Deverbal adjectivalisation

This type may be subject- or object-based. The subject-based deverbal adjectives have the form of lexicalised present participles: *picant* ‘piquant, spicy’, *següent* ‘following’, *vivent* ‘living’; or are derived by a suffix: *eixordador* ‘deafening’, *treballador* ‘hard-working, industrious’, *plover* ‘rainy’, *ploraner* ‘snivelling’, *pensívol* ‘who thinks, meditative, pensive’ (GIEC 2016, 403).

The object-based deverbal adjectives are exclusively suffixations. Their meaning is either active (*agredir* → *agressiu* ‘who aggresses’, *produir* → *productiu*, *venjatiu* ‘who revenges’, *contradir* → *contradictori* ‘that contradicts’, *enganxós* ‘que enganxa’, ‘that hooks’) or passive (*mengívol* ‘that can be eaten, edible’, *admirable* ‘which can be admired’, *creure* → *creïble* ‘credible’, *explicable* ‘explainable’, *oïble* ‘audible’), sometimes including a possibility in the latter case (*espantadís* ‘which can easily be frightened’, *trencadís* ‘which can easily be broken’) (Cabrè 2002, 759–760; GIEC 2016, 401–402).

2.4.2 Relational adjectives

The term “adjectif de relation”, or “relational adjective”, was introduced by Bally (1965, 96–97) and refers to the transposition of a noun to the category of adjective, establishing a relationship between the underlying noun without changing its meaning and the noun which appears in the syntax, e.g. *normes ortogràfiques* ‘normes de l’ortografia’, ‘spelling rules’, a typical equivalent in the Germanic languages that motivates nominal equivalents in English. Note that these adjectives may develop systematic secondary meanings which permit other uses than relational ones, i.e. in *bestial*, *cavalleresc*, etc. (see below). The suffixes *-à -ana*, *-ívol -a*, *-al*, *-enc -a*, *-er -a*, *-ès -a*, *-esc -a*, and *-í -ina* serve to derive adjectives of this type: *ciutat* → *ciutadà* ‘of the town, town(-)’, *valencià* ‘of Valencia, Valencian’, *lul·lià* ‘of Lluïa’, *baró* → *baronívol* ‘man’s, manly’, *pagesívol* ‘of a peasant’, *bestial* ‘bestial’ *bisbal* ‘of a bishop’, *reial* ‘king’s’, *canadenc* ‘Canadian’, *eivissenc* ‘of Eivissa’, *estiuenc* ‘summer-’, *nadalenc*

‘Christmas(-)’, *fronterer* ‘of the frontier’, *mentida* → *mentider* ‘lying, deceitful’, *rialla* → *rialler* ‘jolly’, *francès* ‘French’, *xinès* ‘Chinese’, *cavalleresc* ‘knightly, chivalrous’, *barceloní* ‘of Barcelona’, *boví* ‘cattle, bovine’, *gegantí* ‘gigantic’, *mallorquí* ‘of Majorca’ (Cabré 2002, 757–758). The suffix *-ós, -osa* implies a quantification, for instance in *polsós* ‘full of dust’, *boirós* ‘foggy, misty’, *neguitós* ‘unquiet, restless’, as do *-at, -ut*: *barbat* ‘bearded’, *banyut* ‘horned’, *llanut* ‘woolly’, *panxut* ‘big-bellied’ (GIEC 2016, 402).

The paraphrase of the type *internacional* ‘between nations’, in *comerç internacional* ‘trade between nations’, shows that it is not derived from *inter + nacional*, in which case it would be a prefixation – it seems to be impossible to give an adequate paraphrase of this very common interpretation –, but from the Latinate *internacion- + -al*. Similar cases include *transpirenenc* ‘dellà dels Pirineus’, ‘on the other side of the Pyrenees’, *contranatural* ‘contra la natura’, ‘against nature’, *submarí* ‘sota el mar’, ‘under the sea’, *extraterritorial* ‘fora del territori’, ‘outside the territory’ (Gràcia 2002, 793–794; GIEC 2016, 428, 429–432; cf. 2.3.3).

A subtype of the former, which may be converted to a noun, is the mixed Neo-Latin and native relational adjective formed by means of *anti-* or *post-*, both suppletive forms of the prepositions *contra* and *després de* respectively, for example in *antirovell* in *liquid antirovell* ‘liquid against rust, i.e. rust prevention agent’ or *postvenda* in *servei postvenda* ‘service after sale’. Both prefixes clearly establish a relationship between *liquid* or *servei* (A) and *rovell* or *venda* (B) (Gràcia 2002, 795; cf., however, GIEC 2016, 425).

2.5 Adverbialisation

Regular adverbs are formed using the feminine form of the adjective, which maintains its accent and is considered a compound for this formal reason by Mascaró (1985, 70–71), followed by the suffix *-ment*: *fort, forta* → *fortament*. Sometimes, the masculine form is converted into an adverb, in particular in combination with certain verbs: *parlar clar* ‘to speak clearly’, apart from occasional forms such as *a bocons* ‘on one’s face’ or *de genollons* ‘on one’s knees’ (Rull 2009a; 2009b; GIEC 2016, 407).

3 Modification

As in grammar, derivatives are modifications of the base. The determinant follows the determinatum in the case of suffixation and precedes it in prefixed derivatives. Nouns, adjectives, verbs and adverbs will be treated together when a suffix serves to coin words belonging to these word classes.

3.1 Feminine gender

The feminine gender is marked by the ending *-a* in the designation of persons, for example *germà* ‘brother’ → *germana* ‘sister’, *nebot* ‘nephew’ → *neboda* ‘niece’, *espòs* ‘husband’ → *esposa* ‘wife’, *noi* ‘boy’ → *noia* ‘girl’, for female animals in *lleó* ‘lion’ → *lleona* ‘lioness’, *gos* ‘dog’ → *gossa* ‘bitch’, and by the addition of a suffix, i.e. *-essa* in *mestre* ‘master, teacher’ → *mestressa* ‘lady, landlady, teacher’ or *-ista* without gender distinction in *artista*. Occasionally, *-ot* forms the male counterpart of feminine lexemes (*bruixa* ‘witch, sorcerer’ → *bruixot* ‘wizard, sorcerer’, *abella* ‘bee’ → *abel·lot* ‘drone, bumblebee’, *merla* ‘(female) blackbird’ → *merlot* ‘male blackbird’). We may add *jardiner* → *jardinera* ‘gardener’ as an example of the change of gender in numerous suffixes (GIEC 2016, 412).

3.2 The formation of collective nouns

The formation of collective nouns consists in the singularisation of a plural and is fragmented into blurred referential domains including persons, animals, plants and objects. The suffixes are not neatly specialised, as can be noted in the use of *-ada* for persons and animals (*gentada* ‘crowd (of people)’, *cavallada* ‘herd of horses’, *vacada* ‘herd of cows’, including words such as *boirada* ‘fog bank’), *-alla* for persons, animals and objects (*joves* ‘young persons’ → *jovenalla* ‘young people’, *polls* ‘younglings’ → *pollalla* ‘chicks’, *ferro* ‘iron’ → *ferralla* ‘scrap iron’), *-am* for persons, animals and objects as well (*criaturam* ‘children’, *moscam* ‘swarm of flies’, *branques* ‘boughs, branches’ → *brancam* ‘boughs, branches’), *-ar* for trees and animals (*alzines* ‘holm oaks’ → *alzinar* ‘holm oak forest’, *pins* ‘pines’ → *pinar* ‘pine-forest’, *pomerar* ‘plantation of apple trees’, *bèsties* ‘animals’ → *bestiar* ‘livestock, cattle’), *-atge* for objects and persons (*brancatge* ‘boughs, branches’, *paisanatge* ‘civilian population’) and *-eda* for trees (*faigs* ‘beeches’ → *fageda* ‘beech-forest’).

3.3 Diminutives

The boundaries between diminutives and pejoratives or augmentatives and pejoratives overlap with regard to their semantic classification. The basic diminutive suffix is *-et*, *-eta* and is used for quantification and, whenever diminution cannot apply, for emotive meaning in nouns, adjectives and adverbs such as *casa* ‘house’ → *caseta* ‘little or small house’, *cavallet* ‘little horse, rocking horse’, *homenet* ‘little man, manikin’, *peuet* ‘small foot’, *platet* ‘small plate’; *bonic*, *-a* ‘pretty’ → *boniquet*, *-a* ‘quite pretty’, *grandet*, ‘rather tall, quite grown-up’, *petitet* ‘very small’; *a prop a prop* ‘near at hand’ → *a prop a propet* ‘very near at hand’, *aviat* ‘soon’ → *aviadet* ‘quite soon’, *lluny* ‘a little far’, *a poc a poc* ‘not so fast!, slowly’; and, rarely, verbs in combina-

tion with another suffix (*volar* ‘to fly’ → *voletejar* ‘to flutter’). The suffix *-ó* is used in lexicalised derivatives (*carrer* ‘street’ → *carreró* ‘narrow street or lane’) and in the formation of adjectives (*petitó* ‘very small, very little’). Both suffixes may be combined, i.e. in *petitonet* (GIEC 2016, 409).

The formation of diminutive verbs is idiosyncratic and admits an aspectual, frequentative and emotive interpretation, for example in *ploure* ‘to rain’ → *ploviscar* ‘to drizzle’, *parlar* ‘to speak’ → *parlotejar* ‘to chat, to chatter’, *menjar* ‘to eat’ → *menjotejar*, *menjussar* ‘to pick at one’s food’ (GIEC 2016, 410).

Certain suffixes may be interpreted as denoting approximation, i.e. *-enc*, *-a* in *agre*, *-a* ‘sour’ → *agrenc* ‘slightly sour’, *blau*, *blava* ‘blue’ → *blavenc* ‘bluish’, *dolç*, *-a* ‘sweet’ → *dolcenc* ‘sweetish’; *-ís*, *issa* in *blavís* ‘bluish’, *malalt* ‘sick’ → *malaltís* ‘in poor health’; and *-ós*, *-osa* in *amarg* ‘bitter’ → *amargós* ‘slightly bitter’, *blavós* ‘bluish’, *humit*, *-a* ‘damp’ → *humitós* ‘somewhat damp’ (GIEC 2016, 403).

3.4 Augmentatives

The formation of augmentative nouns and adjectives is less conspicuous than diminution and is expressed currently by *-às*, *-assa* in *home* ‘man’ → *homenàs* ‘big, fat, stout man’, *noi* ‘boy’ → *noiàs* ‘big boy’, *casa* ‘house’ → *casassa* ‘big house’ or *gran* ‘great, big’ → *grandàs* ‘very great, very big’ and by *-ot*, *-a* in *cuixa* ‘thigh’ → *cuixota* ‘big thigh’, *home* ‘man’ → *homenot* ‘big man, fellow, chap’ and *gros*, *grossa* ‘big’ → *grossot* ‘very big’, *mentider*, *-a* ‘mendacious’ → *mentiderot* ‘very mendacious’ (GIEC 2016, 409).

3.5 Pejoratives

Pejorative nouns, adjectives and verbs are formed using the suffix *-ot*, *-ota*, as in lexicalised words such as *cavall* ‘horse’ → *cavallot* ‘nag’, by *-arro*, *-a* in *cotxe* ‘car’ → *cotxarro* ‘big car’, *cuixarra* ‘big thigh’, always used negatively or disrespectfully in the preceding and following examples, by *-atxo*, *-a* + *-ó*, *-ona* in *prim*, *-a* ‘thin, slim’ → *primatxo* ‘rather meagre’ and *-iny-* in *plorar* ‘to cry, to weep’ → *plorinyar* ‘to whimper’.

3.6 Intensification

Intensification appears in nouns denoting time considered as periods of time, i.e. *any* ‘year’ → *anyada*, as ‘whole year’, ‘year of service’, etc. and likewise *matinada* ‘morning’, *mesada* ‘month’, *setmanada* ‘week’.

The elative meaning, also improperly referred to as the superlative, is a form of intensification in adjectives and adverbs which is expressed by *-íssim*, *-a* in *antic*, *-ga*

‘old’ → *antiquíssim* ‘very old’ and *re-* in *sec*, *-a* ‘dry’ → *ressec*, *-a* ‘very dry’, *bé* ‘well’ → *rebé* ‘very well’.

3.7 Negation

The prefixes of negation serve to modify all the lexical word classes. The prefix *in-* derives adjectives (as well as adverbs and deadjectival predicate nouns as in the other cases) in *feliç* ‘happy’ → *infeliç* ‘unhappy’ (*infeliçment*, *infelicitat*), *imprecís* ‘unexact’ (*imprecisament*, *imprecisió*), *il·legible* ‘illegible’ (*il·legibilitat*), *a-*, allomorph *an-* in vowel-initial bases of Greek origin, forms adjectives in *atípic*, *-a* ‘untypical’, *anorgànic*, *-a* ‘inorganic’, occasionally *mal-* in *malcreient* ‘unbelieving’. Apart from adjectives and nouns, *des-* and its learned variant *dis-* form verbs as well, as in *desordre* ‘disorder’, *dissort* ‘misfortune’; *deslleial* ‘unfaithful’ (*deslleialtat*, *deslleialment*), *deshonest* ‘dishonest’ (*deshonestament*, *deshonestedat*), *discontinu* ‘incessant, discontinuous’ (*discontinuitat*), *disconforme* ‘not concurrent’ (*disconformitat*); *desfer* ‘to unmake, to disassemble’, *embolicar* ‘to wrap (up)’ → *desembolicar* ‘to unwrap, to unpack’, *culpar* ‘to blame’ → *disculpar* ‘to excuse’. The reversative function of *des-* and *dis-* is interpreted here in terms of negation (GIEC 2016, 426–428).

3.8 Localisation

Generally, prepositions and adverbs take the form of prefixes in the derivatives of verbs, nouns and adjectives. This semantic type is here called “localisation” on the assumption that the spatial meaning is primary and the temporal and conceptual meanings secondary. Presupposing that prepositions serve to situate or localise two elements, A and B, in space, time and conceptually (cf. 2.3.3), this interpretation also applies to types such as *contraquerella* ‘countercharge’, which is ‘(a charge (A)) against another charge (B)’, or *contraverí* ‘antidote’, ‘(a substance (A)) against a poison (B)’, and *sensepapers* ‘(a person (A)) without papers’. This explanation allows for the fact that A and B may have the same lexical meaning (*querella* in *contraquerella*) or a different one (such as ‘substance (A)’ in *contraverí*, since the ‘substance’ is not necessarily a poison). Evidently, the cohesion between the prepositional prefix and B is stronger than the one between the covert A and the rest, and both elements must be identified in the context. The forms, mostly learned, and their meanings are lexicalised to differing degrees. Some meanings can be ordered in semantic terms, for example ‘before’ (*avant-*, *ante-*, *pre-*) and ‘after’ (*rere-*, *post-*): *avantbraç* ‘forearm’ (Mascaró 1985, 64, 70), *antesala* ‘room before a room’, *antepenúltim* ‘before the penultimate’, *antedatar* ‘to date before’, *prehistòria* ‘time before history’, *preveure* ‘to foresee’, *postpart* ‘period after birth’, (*servei*) *postvenda* ‘(service) after sale’, *postdatar* ‘to date after (later)’. Likewise, the meaning of *co(n)-* ‘with’ (Cat. *amb*) is compatible

with nouns, adjectives and verbs as in *coautor* ‘co-author’, *coexistir* ‘to coexist’, *convivre* ‘to live together with’. Other prefixes appear in *sotabosc* ‘plant(s) under the forest, undergrowth’, *superhome* ‘superman’, *sobreexcitar* ‘to overexcite’, *entreobrir* ‘to open half’, etc. (Mascaró 1985, 70; Gràcia 2002, 787–793; GIEC 2016, 428–429).

However, some prefixes do not belong to this class but express a position in a hierarchy, as *arque-/arxi-*, *bes-*, *re-* and *vice-* show when added to a noun: *arquebisbe* ‘archbishop’, *arxiduc* ‘archduke’, *besavi* ‘great-grandfather’, *rebesavi* ‘great-great-grandfather’, *vicepresident*.

3.9 Lexical aspect

Lexical aspect can be found in the form of the repetitive meaning ‘again’ (*reelegir* ‘to re-elect’) or the durative meaning ‘to continue (doing)’ (*revendre* ‘to resell’) and the frequentative meaning expressed by *-ejar* (*voletejar*, *desvariar* ‘to fantasize’ → *desvariejar* ‘to fantasize’).

3.10 Formation of transitive (causative) and reflexive verbs

A small group of intransitive verbs is modified by means of the prefix *a-* and forms transitive verbs with causative meaning: *jeure* ‘to lie’ → *ajeure* ‘to put, to lay’, *dormir* ‘to sleep’ → *adormir* ‘to lull sb to sleep’, *seure* ‘to sit’ → *asseure* ‘to put, to place’, and may be used as reflexive verbs as well: *ajeure*’s, *adormir-se*, *asseure*’s.

4 Compounding

4.1 Deverbal generic compounding

What is called “generic compounding” (Staib 1988) in this semantic approach, is considered under the morphological category of derivation in Cabré (2002); Rull (2004, 221–254); GIEC (2016, 396–401).

4.1.1 Verb + generic element

In the compounds formed by verb + the generic element ‘somebody’, the latter is the determinatum and the verb the determinant. The person has the function of subject and, accordingly, the suffixes serve to form subject nominalisations. The suffixes *-dor*, *-a* and *-aire* are semantically neutral (cf. Martines 1997 on *-aire* in Valencian), whereas the derivatives formed by *-ant/-ent* are similar to the present participle by the fact that

the suffix refers to an action presented as being in progress and having a certain duration. The meaning can be paraphrased by the formula ‘sb who hunts’ in *caçador* ‘hunter’ and in *cosidora* ‘seamstress’, *treballador* ‘worker’; *dansaire* ‘dancer’, *xerraire* ‘talker, chatterbox’; *caminant* ‘hiker, wayfarer’, *estudiant* ‘student’, *aprenent* ‘apprentice’ (Grossmann 1998a; 1998b; GIEC 2016, 400–401).

The generic element may denote ‘something’ that has the function of an instrument or a place and exhibits varying degrees of lexicalisation that are ignored in the paraphrase of the type ‘sth with which someone scares sb’ in *espantall* ‘scarecrow’ and in *fermall* ‘fastener’, *mirall* ‘(looking-)glass’, *mocador* ‘handkerchief’, *penjador* ‘(coat-)hanger’, *regadora* ‘watering-can’, *sonall* ‘little bell’, *tapadora* ‘lid’ or ‘place where sb dines’ in *menjador* ‘dining-room’, *escorxador* ‘slaughter-house’, *rebedor* ‘hall, entrance room’. The learned suffix *-tori* is synchronically analysable in *dormitori* and *observatori* (Grossmann 1998a; 1998b; GIEC 2016, 400).

4.1.2 Denominal relational compounding

The above term is proposed here in accordance with the motivation of the term “relational adjective” (2.4.2). As far as the pattern is concerned, it is possible to ascertain only a relationship between a noun and ‘sb’ or ‘sth’ as in *ferro* ‘iron’ + ‘sb’ → *ferrer* ‘blacksmith, smith’ and *formiga* ‘ant’ + ‘something’ → *formiguer* ‘anthill’. Generally, the lexical meaning, which contains a verb in the lexical definition as in the case of deverbal generic compounding, is taken for granted as the pattern meaning, i.e. ‘person who works iron’ and ‘hill/nest where ants live in society’, but a speaker who only knows *ferro* ‘iron’ and *formiga* ‘ant’ cannot deduce the exact lexical meaning from this information nor a particular verb provided by his or her general knowledge rather than any linguistic competence.

The compounds denoting person correspond to the type noun + ‘sb’ in suffixations by means of *-er*, *-a*, *-aire* and learned *-ista*: *ferrer* ‘blacksmith, smith’, *fuster* ‘joiner’, *porter*, *-a* ‘porter’; *terrisa* ‘pottery’ → *terrisaire* ‘potter’; *artista*, *dentista*.

Identical in part to those of the former type, the suffixes used to derive the type noun + ‘sth’ are applied to trees and objects and *-era* to other plants as well as objects, *-al* to objects, *-eria* and *-ar* to places, a possible overlapping with the formation of collective nouns (4.2): *albercoquer* ‘apricot tree’, *ametller* ‘almond tree’, *cirerer* ‘cherry tree’, *cendrer* ‘ashtray’; *alfabreguera* ‘basil’, *favera* ‘broad bean (plant)’, *maduixera* ‘strawberry (plant)’, *suçrera* ‘sugar basin’; *didal* ‘thimble’; *rellotgeria* ‘watchmaker’s shop’, *sastreria* ‘tailor’s business’; *colomar* ‘pigeonry, dovecote’.

4.2 Lexeme compounding

There are two main approaches in the treatment of Catalan composition, which concern the position of prefixation in word-formation and the limits of compounding and syntax. Catalan prefixes are split up into two groups on formal grounds. The usual pattern is either an unstressed (*a-*, *en-*, *in-*, etc.) or a stressed (*contra-*, *entre-*, *sobre-*, etc.) form. Cabré (2002, 734–736) assigns the unstressed prefixes to derivation and the stressed prefixes to compounding (Gràcia 2002, 785–796) based exclusively on formal criteria in the tradition of 19th century linguistics (Lüdtke 2011, 202–203). It goes without saying that this is one of the major differences between a description in morphological or semantic terms (cf. 2.3.3, 2.3.6, 2.4.2, 3.7–3.10).

“For a combination to be a compound one condition has to be fulfilled: the compound must be morphologically isolated from a parallel syntactic group” (Marchand 1969, 21). This criterion marks the boundary between compounding and phraseology (75.5 Lexicalized Syntax: Phraseology). Usually, combinations which are constructed according to grammatical rules, currently called ‘syntagmatic compounds’ in Catalan linguistics, are considered compounds when they are lexicalised and manifest various degrees of grammatical integration (GIEC 2016, 447–448, 456–457). However, words such as *allioli* ‘aioli’, *aiguardent* ‘strong liquor’, *celobert* ‘air shaft’ are formed as *all i oli*, *aigua ardent* and *cel obert*. Likewise, *cap-gros* ‘big head, tadpole’, *pit-roig*, *pell-roja*, *centcomes*, *set-ciències* are cited as compounds because they are interpreted as exocentric compounds in, for instance, *pell-roja* ‘person having a red skin’, *pit-roig* ‘bird having a red breast’. These words are metonyms that are no different from *Verdaguer* in *llegir Verdaguer* ‘to read Verdaguer, i.e. a poem of V.’ and should be treated as such even in the case of lexicalisations. Likewise, lexicalisations of grammatical processes, for example in *daltabaix* ‘(from above) down’, *col-i-flor* ‘cabbage and flower, cauliflower’, will not be included here. If a case like *va i ve* ‘it goes and comes’ → *vaivé* ‘to-ing and fro-ing’ appears here, it is due to the fact that this lexicalised sentence is a conversion (2.1).

Lexeme compounding combines lexemes that may be determined by grammatical processes (*catalanoparlant*, *primmirat*, *somiatruites*) or affixes (*horticultura*). The determinatum either precedes (*catalanoparlant*) or follows the determinant (*filferro* ‘wire’) (cf. GIEC 2016, 449–450). However, some compounds, for example *agredolç* ‘bittersweet’, are of the copulative type. The compounds may be formed on a native basis of coining (*filferro*), on a Neo-Latin (*clarivalent* ‘clear-sighted’) or foreign, generally Germanic, basis (*ferrocarril* according to the pattern of *railway*).

4.2.1 Nouns

In noun + noun compounds the first noun is the head of the construction in *filferro* ‘fil de ferro, string of iron, wire’, *vagó llit* ‘vagó amb llits, railway carriage with beds,

sleeping car', although in rare cases this can also be the second element in *autopista* 'motorway', *mareperla* 'mother of pearl'. Neither element is the head of the construction in *aiguaneu* 'water and snow, sleet' (Gràcia 2002, 805–808; GIEC 2016, 455–456).

Verb-complement compounds, i.e. of the type *parallamps* 'lightning conductor' in Catalan linguistics, is formed by a verb lexeme, sometimes called the third person singular of the present tense (Mascaró 1985, 58–59) but which I prefer to call neutral, followed by a noun in singular or plural. Since the interpretation of this Romance type is controversial, it has to be approached from the act of coining. Its long history in Greek, Latin and Romance makes it clear that it forms adjectives which may be nominalised by conversion or ellipsis, a function still recognisable in the paraphrases of these compounds that are cited in the literature (Gràcia 2002, 801) for a very different purpose, i.e. the secondary use as an adjective in *un noi molt llepaculs* 'a very arse-licking boy' and *una màquina llevaneu* 'a machine which removes snow'. In fact, a similar word, *llepafils* 'who licks threads, choosy about food', is codified as both an adjective and a noun (DGLC). This use is a good indication of how to interpret this type of compound. The problem seems to be the evidence on which the analysis is usually based, i.e. the exclusive use of dictionaries. Research in several Romance languages suggests that neologisms formed using this process are perfunctorily introduced into a language and registered in a dictionary only when their usage is established. Once established, new nouns may be coined directly by conversion. The semantic dispersion of this type is more pronounced than in other processes and is easily accounted for by its adjectival origin; nouns refer to occupations (*guardabosc* 'man who looks after the forest, forest warden'), animals (*trencanous* 'a bird which cracks nuts, hawfinch'), plants (*escanyagats* 'bush which strangles cats, sloe'), instruments (*portallàntia* 'object which holds a candle, candlestick'), objects (*eixugamà*, *-mans* 'object with which sb dries his or her hands') and facetious terms (*menjacapellans* 'man who eats clerics', *escanyamarits* 'widow who strangles husbands, widow in her third marriage', *sanaporcs*, *-truges* 'whistle which serves to announce the castrator who cures pigs/sows', *somiatrutes* 'person who dreams about omelettes, dreamer'). It is not infrequent for the same form to denote different entities, for example in *trencanous* which can refer to a bird or an instrument 'which cracks nuts', although the compounds are mostly formed from recurrent verbal elements such as *cobre-* 'cover', *guarda-* 'look after', *lleva-* 'remove', *pica-* 'prick', *porta-* 'hold', *tapa-* 'cover', *tira-* 'pull', *trenca-* 'break', etc. Considered as either conversion or ellipsis, the semantics seems to favour interpretation as an exocentric compound. If it is necessary to identify a morpheme, the nearest equivalent would be *-er* in *picapedrer* 'man who carves (sth in) stone, stonemason' (for more details see Lüdtke 2011, 386–420; Mascaró 1985, 85–64; Grossmann 1986; Gràcia 2002, 797–802; GIEC 2016, 447, 452–454).

4.2.2 Adjectives

In *agredolç* ‘bittersweet’, *blauverd* ‘blue-green’, *clarobscur* ‘bright and dark, dim’, *sord-mut* ‘deaf and dumb’ (Mascaró 1985, 71–73), the two elements are on the same level; neither is the determinatum or the determinant; thus, they may rightly be called copulative compounds (cf., however, Gràcia 2002, 816–818; *GIEC* 2016, 460–461).

The nouns that are incorporated into the type noun + adjective designate an inalienable part: *bocafi* ‘fi de boca’, *camallarg* ‘llarg de cames, long-legged’, *camacurt* ‘curt de cames, short-legged’, *llenguallarg* ‘llarg de llengua, malicious’, *capalt* ‘amb el cap alt, with the head high, high-spirited’ *capbaix* ‘amb el cap baix, with the head low, low-spirited’, *barbaroig* ‘roig de barba, red with regard to the beard’ (Mascaró 1985, 64–66; Gràcia 2002, 811–814; *GIEC* 2016, 458–459).

The counterpart to the preceding type is the pattern adjective + noun in *verd botella* ‘green like bottle-glass, bottle-green’.

Filiforme ‘in the form of a thread’ and *antropomorf* (Mascaró 1985, 71) are examples of classical compounds.

4.2.3 The type noun + verb or participle

This type incorporates a direct object or an adjunct, generally a part of the body, in a compound: *capgirar* ‘to turn sth on its head, to turn over’, *colltorçar* ‘to bend the neck’, *corprendre* ‘to captivate the heart, to charm’, *ullprendre* ‘to take by the eyes, to bewitch’. Essentially the same structure appears in the type adverb + verb: *carvendre* ‘to sell dearly’, *cartenir* ‘to esteem highly’, *primfilar* ‘to spin finely, to split hairs’, *menysprear* ‘to esteem less, to have a low opinion of’, *maltractar* ‘to ill-treat’ (Mascaró 1985, 66–69; Gràcia 2002, 808–811; *GIEC* 2016, 458, 461–462).

Both types form adjectives that are lexicalised present and past participles on the basis of systematically possible, but undocumented verbs: *bocabadant* ‘opening his or her mouth, voracious’, *bocabadat* ‘whose mouth is open’, *primmirat* ‘who looks meticulously’ (Gràcia 2002, 811–816; *GIEC* 2016, 458–460).

Complement-verb compounds of this type are similar to verb-complement-compounds (4.2.1), but this type is fairly unproductive and primarily forms adjectives and, secondarily, nouns using an inverted order which has consequences for its marking as a present participle: *fefaent* ‘making trust, credible’, or the conversion *terratinent* ‘person owning landed property, landowner’ (Gràcia 2002, 802–804; *GIEC* 2016, 454).

4.3 Reduplicative compounds

This type produces mostly onomatopoeia, for example *xiu-xiu* ‘whisper’, *xim-xim* ‘drizzle’, or other repetitions of one or two syllables with vocalic variation as in

barrim-barram ‘in a hurry’, or without vocalic variation, as shown by *fofo* ‘spongy, wobbly’. These compounds may be nouns or occasionally adjectives and belong to an informal register (Cabr  Monn  2002).

5 Truncation and other formal types

Word-formation goes beyond formal and semantic patterns to create new forms for already existing complex expressions. The most common types are truncation or clipping (conserving the initial element in *independentista* → *indepe* or the final element in *Bartomeu* ‘Bartholomew’ → *Tomeu*), coining by acronyms (*Organitzaci  de les Nacions Unides* → *ONU* ‘UNO’) or blending (*tramvia blau* → *trambiau* ‘blue tramway’) (Cabr  Monn  2002; F bregas Alegret 2014), and will not be treated here.

6 Bibliography

- Badia Margarit, Antonio M. (1962), *Gram tica catalana*, 2 vol., Madrid, Gredos.
- Bally, Charles (?1965), *Linguistique g n rale et linguistique fran aise*, Bern, Francke.
- Bauer, Laurie (1983), *English Word-Formation*, Cambridge, Cambridge University Press.
- Bernhard, Orion (s.a.), *La formaci n de nombres por sufijos en catal n*, manuscript of a thesis without pagination, Zentralbibliothek Z rich; partial publication, Diss. Z rich [1943].
- Blasco Ferrer, Eduardo (1984), *Grammatica storica del catalano*, T bingen, Niemeyer.
- Bruguera i Talleda, Jordi (2006), *Diccionari de la formaci  de mots*, Barcelona, Enciclop dia Catalana.
- Cabr , M. Teresa (1994), *A l’entorn de la paraula*, 2 vol., Val ncia, Universitat de Val ncia.
- Cabr , M. Teresa (2002), *La derivaci *, in: Joan Sol  et al. (edd.), *Gram tica del catal  contempor ni*, vol. 1, Barcelona, Emp ries, 731–774.
- Cabr , M. Teresa/Rigau, Gemma (1986), *Lexicologia i sem ntica*, Barcelona, Enciclop dia Catalana.
- Cabr  Monn , Teresa (2002), *Altres sistemes de formaci  de mots*, in: Joan Sol  et al. (edd.), *Gram tica del catal  contempor ni*, vol. 1, Barcelona, Emp ries, 889–932.
- Coseriu, Eugenio (1978), *La formaci n de palabras desde el punto de vista del contenido (a prop sito del tipo “coupe-papier”)*, in: Eugenio Coseriu, *Gram tica, sem ntica, universales. Estudios de ling  stica funcional*, Madrid, Gredos, 239–264.
- Darmest ter, Ars ne (1875), *Traiti  de la formation des mots compos s dans la langue fran aise*, Paris, Franck.
- DGLC = Fabra, Pompeu (?1974, ?1932), *Diccionari general de la llengua catalana*, Barcelona, L pez Llaus s.
- DIEC = Institut d’Estudis Catalans (?2007), *Diccionari de la llengua catalana*, Barcelona, Edicions 62/ Enciclop dia Catalana.
- Estop , Rosa (2010), *La composici  patrimonial en catal  perd representativitat. Estudi d’un corpus de neologismes de premsa i r dio*, *Estudis Rom nics* 32, 125–147.
- Fabra, Pompeu (1956), *Gram tica catalana*, Barcelona, Teide.
- F bregas Alegret, Immaculada (2014), *La fabrication des n ologismes. La troncation en espagnol et en catalan*, Rennes, Presses Universitaires de Rennes.
- GIEC = Institut d’Estudis Catalans (2016), *Gram tica de la llengua catalana*, Barcelona, Institut d’Estudis Catalans.

- Gràcia, Lluïsa (1995), *Morfologia lèxica. L'herència de l'estructura argumental*, València, Universitat de València.
- Gràcia, Lluïsa (2002), *Formació de mots: composició*, in: Joan Solà et al. (edd.), *Gramàtica del català contemporani*, vol. 1, Barcelona, Empúries, 777–829.
- Grossmann, Maria (1986), *Anàlisi dels compostos catalans del tipus "somiatriutes"*, in: *Miscel·lània Antoni Badia i Margarit*, vol. 4, Barcelona, Publicacions de l'Abadia de Montserrat, 155–169.
- Grossmann, Maria (1989), *La formazione delle parole in catalano. Presentazione di una ricerca in corso con alcune osservazioni sui tipi di verbalizzazione denominale mediante suffissazione*, in: Dieter Kremer (ed.), *Actes du XVIII^e Congrès International de Linguistique et Philologies Romanes*, vol. 7, Tübingen, Niemeyer, 503–511.
- Grossmann, Maria (1994), *Opposizioni direzionali e prefissazione. Analisi morfologica e semantica di verbi egressivi prefissati con "des-" e "es-" in catalano*, Padova, Unipress.
- Grossmann, Maria (1998a), *Formazione dei nomi di agente, strumento e luogo in catalano*, in: Giovanni Ruffino (ed.), *Atti del XXI Congresso Internazionale di Linguistica e Filologia Romanza*, vol. 2, Tübingen, Niemeyer, 383–392.
- Grossmann, Maria (1998b), *Més sobre la formació dels "nomina agentis", "instrumenti" i "loci" en català*, in: Josep Massot i Muntaner (ed.), *Estudis de llengua i literatura en honor de Joan Veny*, vol. 2, Barcelona, Publicacions de l'Abadia de Montserrat, 559–575.
- Lang, Jürgen (1991), *Die französischen Präpositionen. Funktion und Bedeutung*, Heidelberg, Winter.
- López del Castillo, Lluís (2002), *Diccionari de formació de paraules*, Barcelona, Edicions 62.
- Lüdtke, Jens (1978), *Prädikative Nominalisierungen mit Suffixen im Französischen, Katalanischen und Spanischen*, Tübingen, Niemeyer.
- Lüdtke, Jens (1983), *L'originalitat de la nominalització catalana*, *Estudis Universitaris Catalans* 25, 331–335 (*Miscel·lània Aramon i Serra. Estudis de llengua i literatura catalanes oferts a R. Aramon i Serra en el seu setantè aniversari III*).
- Lüdtke, Jens (2011), *La formación de palabras en las lenguas románicas: su semántica en diacronía y sincronía*, Mexico D. F., El Colegio de México.
- Marchand, Hans (1969), *The Categories and Types of Present-Day English Word-Formation*, München, Beck.
- Martines, Josep (1997), *El sufix "-aire" al País Valencià*, in: Josep Massot i Muntaner (ed.), *Miscel·lània Germà Colon*, vol. 7, *Estudis de Llengua i Literatura Catalanes* 34, 229–262.
- Mascaró, Joan (1985), *Morfologia*, Barcelona, Enciclopèdia Catalana.
- Moll, Ana (1957), *Sufijos nominales y adjetivales en ibicenco*, *Revista de Filología Española* 41, 341–371.
- Moll, Francesc de Borja (1968), *Gramàtica catalana especialment referida a les Illes Balears*, Palma, Moll.
- Rull, Xavier (2004), *La formació de mots. Qüestions de normativa*, Vic, Eumo.
- Rull, Xavier (2009a), *La composició culta en català*, Palma de Mallorca, Moll.
- Rull, Xavier (2009b), *De la sufixació en català. Apunts i reflexions (1999–2009)*, Benicarló, Onada Edicions.
- Rull, Xavier (2016), *The Behavior of Prefixed, Suffixed and Compound "Nomina Actionis" in Catalan with Regard to Conversion: a Sign of Underspecification?*, Aachen, Shaker.
- Solà, Joan, et al. (edd.) (2002), *Gramàtica del català contemporani*, vol. 1, Barcelona, Empúries.
- Staib, Bruno (1988), *Generische Komposita. Funktionelle Untersuchungen zum Französischen und Spanischen*, Tübingen, Niemeyer.
- Weidenbusch, Waltraud (1993), *Funktionen der Präfigierung. Präpositionale Elemente in der Wortbildung des Französischen*, Tübingen, Niemeyer.

8 Variation and Varieties

Mar Massanell i Messalles

8.1 Dialects

Abstract: This chapter on Catalan dialects contains three main sections followed by some concluding remarks. The first section deals with the linguistic area of Catalan, which is unevenly spread across four modern European states: Spain, Andorra, France and Italy. The second section presents the dialectal divisions of Catalan and consists of two subsections. The first subsection discusses the broad divide between Western and Eastern Catalan, which is based on the linguistic criteria proposed by Milà i Fontanals (1861), Fabra (1906) and Veny (¹²1998, ¹1978). The second subsection, essentially following Alcover (1926) but rounded out by the subsequent work of Veny (¹²1998, ¹1978), describes the further division of these two main blocks into six dialects, with Western Catalan subdivided into Northwestern and Valencian, and Eastern Catalan subdivided into Northern or Rossellonese, Central, Balearic and Alguerese. The third section provides further details on the main characteristics of these six dialects. The chapter closes with a summary of the fundamental linguistic variables affecting the dialectal diversity of Catalan and some final observations.

Keywords: linguistic domain, Western-Eastern division, Northwestern Catalan, Valencian, Northern Catalan or Rossellonese, Central Catalan, Balearic, Alguerese

1 The linguistic area

Catalan is a language derived from Vulgar Latin that developed at some point before the 9th century north and south of the Eastern Pyrenees (↗10 Early Medieval Catalan), in territories that are now administratively divided between the French *département* of Pyrénées Orientales, the Principality of Andorra and the Spanish autonomous communities of Catalonia and Aragon. Thereafter, beginning in the 11th century and at an accelerated pace in the 13th century, Catalan was spread through conquest and repopulation to lands further south and east along the Mediterranean coast and to nearby islands (↗11 The Growth and Expansion of Catalan (1213–1516)). It is this history that partially accounts for the dialectal differences that characterise the various areas where Catalan is currently spoken.

Today, the linguistic area is unevenly spread across four European states – Spain, Andorra, France and Italy – in which both the sociolinguistic position of

<https://doi.org/10.1515/9783110450408-014>

the language (↗18 Language Demography) and its legal status (↗19 Language Law and Language Policies) vary considerably. The largest part of the linguistic domain is found within Spain, where Catalan is spoken in the five autonomous communities of Catalonia, Aragon, the Balearic Islands, Valencia and Murcia. The first and foremost of these is Catalonia, which is divided into 41 regions called *comarques* (sing. *comarca*) and one singular territorial entity, Aran, which falls within the language domain of Occitan rather than Catalan. In the neighbouring community of Aragon, Catalan is spoken in a strip of land adjacent to Catalonia, popularly known as *la Franja de Ponent*, which comprises the four regions of Ribagorça, Llitera, Baix Cinca and Matarranya (Monclús i Esteban 2014). Catalan is also spoken in the autonomous community of the Balearic Islands, comprised of Majorca, Minorca and the Pityusic Islands (Eivissa and Formentera), which were fully repopulated by Catalans after James I the Conqueror defeated Muslim rulers in Majorca (1229) and Eivissa and Formentera (1235) while Minorca fell to Alfons the Frank (1287). At around the same time, James I defeated the Almohads in most of what is now the autonomous community of Valencia, leading to the flight of much of the local Muslim population. The settlers who replaced them were largely Catalan-speaking, with the exception of a few regions in the northern interior (Racó d'Ademús, Alt Millars, Alt Palància and Serrans), where Aragonese-speaking settlers predominated (Guinot Rodríguez 1999). However, the historical predominance of Catalan in the Valencian Country was slowly reduced from the 17th century, since some of the lands made available by the expulsion of the Moriscos (1609–1614) were repopulated with Spanish-speaking settlers (Foia de Bunyol, Canal de Navarrés and Vall de Cofrents), and epidemics in 1648 also contributed to the introduction of Spanish to the region of Baix Segura (Casanova 2001). Moreover, in the 19th century, the official boundaries of the Valencian Country were altered to include the Marquisate of Villena and the area of Plana d'Utiel, whose population had never been historically Catalan-speaking, resulting in roughly 40 % of the area of the Valencian Country being non-Catalan-speaking (Casanova 2001). Finally, Catalan is spoken in a tiny region of Murcia called El Carxe. This Catalan-speaking area consists of small villages belonging to three municipalities (Iecla, Jumella and Favarella) and its existence is the result of farmers from the neighbouring Valencian Vinalopó valley repopulating the empty lands in the 19th century (Limorti/Quintana 2012).

Catalan is also the language spoken in the Principality of Andorra (Molla 2003), a sovereign microstate nestled between Spain and France in the Eastern Pyrenees. Within France itself, Catalan is historically the local language in the regions of Rosselló, Vallespir, Conflent, Capcir and Alta Cerdanya, which together constitute what is known as *Northern Catalonia* (Baylac-Ferrer 2009). These areas were severed from the rest of Catalonia and annexed by France under the Treaty of the Pyrenees signed in 1659 by Philip IV of Spain and Louis XIV of France (↗12 The Origins of Modern Catalan: Cultural and Linguistic Evolution). Currently, together with the Occitan region of Fenolleda, these lands constitute the *Département des Pyrénées*

Orientales.¹ Finally, Catalan is still spoken in the town of l'Alguer, a linguistic enclave on the northwest coast of the Italian island of Sardinia (Corbera/Domènech/Mayoral 2000; Bover i Font 2002). It was introduced there following the Crown of Aragon's conquest of the island in 1354 under Peter the Ceremonious, when, after repeated revolts, the inhabitants of the town were expelled and replaced with Catalan settlers in 1372.

2 The dialectal divisions of Catalan

Catalan is divided into two broad dialect blocks, the more conservative Western Catalan and the more innovative Eastern Catalan. The two areas are most easily distinguished by their respective pronunciation of the vowels *a* and *e* in unstressed syllables. On one side of the isogloss, in Western Catalan, the vowels maintain their full values (e.g., *casa* ['kaza] 'house', *pare* ['pare] 'father'). On the other side, in Eastern Catalan, they are conflated in a schwa (e.g., *casa* ['kazə] 'house', *pare* ['parə] 'father'). Several other isoglosses pass close to this main one, together forming an isogloss bundle that clearly demarcates this overarching dialectal division. Both of these main branches of Catalan can be further subdivided into dialects on the basis of the first person present tense ending for first conjugation verbs. For instance, for the verb *cantar* 'to sing', the first person singular present tense 'I sing' is *cant* in Balearic and Alguerese, *cante* in Valencian, *canti* in Northern Catalan and *canto* in both Northwestern and Central Catalan, but with the *-o* pronounced [o] and [u] respectively. Western Catalan comprises two dialects, Northwestern Catalan and Valencian, while Eastern Catalan is made up of four, Northern Catalan or Rossellonese, Central Catalan, Balearic and Alguerese.

The differences between these two dialect blocks are essentially due to linguistic innovations that occurred in Eastern Catalonia (where the capital city of Barcelona is located) starting as early as the 13th century (Recasens 2017, 411–424). The dialects spoken in territories where the presence of Catalan is due not to the evolution of a locally spoken form of Vulgar Latin but rather to the influx of a Catalan-speaking population (named *consecutive dialects* as opposed to *constitutive dialects*) are situated in either one dialect block or the other depending on the geographical origin of the majority of the Catalan settlers in that particular area (Griera 1931, 3–17; Ferrando Francés 1989). In the case of the Valencian Country, in addition, the presence of the minority Aragonese-speaking settlers has played a role in strengthening certain linguistic features that this language shares with Western Catalan (Alarcos Llorach

¹ As an interesting aside, while it is well established that the gypsy community of Perpinyà has played an important role in the survival of Catalan in Northern Catalonia (Escudero 2004), it has recently also been shown that Catalan-speaking gypsy groups scattered throughout France have spread their language well beyond the historical limits of Catalan (Casanova Solanes 2016).

1960; Ferrando Francés 1989; Ponsoda Sanmartín 1996), and recent studies have shown that many Valencian words thus far regarded as of Mozarabic origin are in fact derived from Aragonese (Martines 2002, 160-161; 2009).

2.1 The Western-Eastern divide

This basic division of Catalan runs north-south through Catalonia, with the Catalan spoken in the western part of the community joining the varieties spoken in Andorra, the Aragonese Strip, the Valencian Country and El Carxe in Murcia to comprise Western Catalan, while the Catalan of Eastern Catalonia together with the dialects of Northern Catalonia in France, the Balearic Islands and l'Alguer in Sardinia comprise Eastern Catalan. In Catalonia, the Western-Eastern division follows natural geographic boundaries determined by topography; Western Catalan is spoken in the lands comprising the watershed of the rivers Segre and Ebre and their tributaries, while Eastern Catalan is spoken in the river basins of the Fluvià, Ter, Tordera, Besòs, Llobregat, Foix, Gaià, Francolí, Riudoms, Alforja and Riudecanyes. Because the administrative division of Catalonia into *comarques* is also geographically based, dialect boundaries tend to coincide with administrative regional boundaries, the chief exception being the Solsonès region, which is bisected by the main dialectal isoglosses.

2.1.1 The basis for the Western-Eastern divide as proposed by Milà i Fontanals (1861)

The division of Catalan varieties between Western and Eastern was first postulated by Milà i Fontanals (1861, 462) on the basis of what he observed with regard to unstressed vowels. In Catalan, while there are seven vowels that can appear in stressed syllables (*mà* ['ma] 'hand', *pèl* ['pɛl] 'hair', *vent* ['ben] 'wind', *fi* ['fi] 'end', *dol* ['dɔl] 'grief', *molt* ['mol] 'much', *suc* ['suk] 'juice'),² in unstressed syllables these vowels are reduced to five in Western Catalan (as it was in what is known as *Preliterary Catalan* from the 9th to the 12th centuries) and three in Eastern Catalan as a result of a shift that began to take place in the 13th century in palatal and central vowels (Coromines 1971, 295–296) and at the end of the 15th century in velar vowels (Veny ¹²1998, 30–31). The resulting Western-Eastern distinction can be seen in the respective pronunciations of derived words like *manat* [ma'nat]/[mə'nat] 'handful', *pelut* [pe'lut]/[pə'lut] 'hairy', *ventós* [ben'tos]/[bən'tos] 'windy', *dolgut* [dɔl'ɣut]/[dul'ɣut] 'grieved' and *moltíssim* [mol'tisim]/[mul'tisim] 'very much'. By contrast, the two dialects share their pronunciation of words like *final* [fi'nal] 'ending' and *sucós* [su'kos] 'juicy'.

² Some Balearic varieties have one more vowel (/ə/), while Northern Catalan has a reduced vowel inventory (/i/, /e/, /a/, /ɔ/, /u/), as we will see in 3.2.

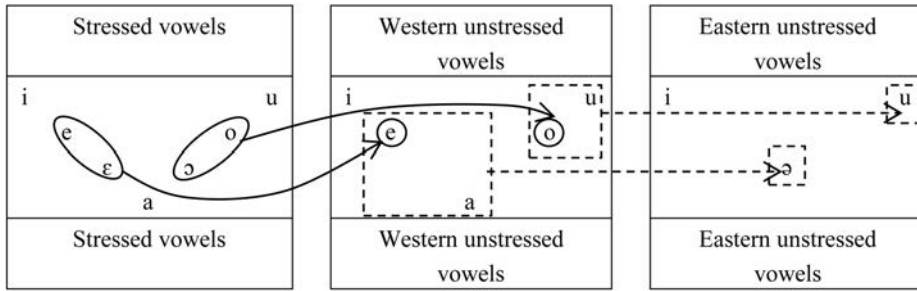


Figure 1: Vowels in stressed syllables in Catalan, and the same vowels as they are pronounced in unstressed syllables in Western and Eastern Catalan.

However, this description of the behaviour of Catalan unstressed vowels requires some further specification. For example, with regard to Eastern Catalan (Recasens 1990–1991, 284), Catalan as spoken in Majorca seems to represent an intermediate linguistic stage between the more conservative Western Catalan on the one hand and all other Eastern Catalan dialects on the other, given that in Majorcan, the unstressed *a* and *e* have converged in [ə] (as in Eastern Catalan), but the unstressed *o* and *u* are kept distinct in most linguistic contexts (as in Western Catalan). And in l'Alguer, the historical [ə] found in the speech of the 14th century Catalan settlers³ has been opened to [a] by contact with Sardinian, which lacks that vowel.

On the other hand, the situation with regard to Western Catalan is likewise not so straightforward. In fact, post-tonic *a* in final position may take different pronunciations. Thus, in the Catalan regions of Segrià, Noguera, Pla d'Urgell, Urgell and Garrigues, this vowel is pronounced [ɛ]: *padrina* [pa'ðrine] 'godmother; grandmother',⁴ *pedra* ['peðre] 'stone', (*jo*) *cantava* [kan'taβɛ], [kan'taɛ] or [kan'taje] '(I) sang' (Gili i Gaya 1931, 9–11; Solans 1996, 34). By contrast, in Southern Valencian the final unstressed *a* is affected by a phenomenon of vowel harmony which can take place, depending on the particular locality, when it is preceded in the stressed syllable by an open *e*, as in *terra* ['tere] 'land' and the place name *Sueca* [su'eke], and/or by an open *o*, as in *dona* ['dɔnɔ] 'woman' and *olla* ['ɔʎɔ] 'pot', and in some places the pronunciation of *a* in final position has become generalised to either [ɛ] or [ɔ] (Hadwiger 1905; Colomina i Cas-

³ It seems that the Catalan settlers of l'Alguer came mostly from other parts of the Eastern Catalan-speaking area. According to Conde y Delgado de Molina (1994, 92), we know the geographical origin of 21 of the first Catalan settlers: seven came from Barcelona, two from Tarragona, one from Vilafranca del Penedès, one from Cervià de Ter (all of these in Eastern Catalonia); two from Perpinyà and one from Cotlliure (in Northern Catalonia), and four from Majorca; and only three from Valencia, in the Western Catalan-speaking area.

⁴ The original meaning of *padri* is 'godfather', but the custom by which the godfather is generally the grandfather has caused this semantic shift in Northwestern Catalan, Northern Catalan and Majorcan (Veny 2007–2017, vol. 3, map 364).

tanyer 1985, 75–80). Moreover, in many Western varieties, in a pretonic syllable *e* may open to become [a], above all at the beginning of a word and especially if the *e* is followed by a sibilant or nasal, as seen in *escola* [as'kola] 'school', *eixut* [aj'jut] 'dry', *enciam* [ansi'am] 'lettuce', and *embut* [am'but] 'funnel', but also *eriçó* [ari'so] 'hedgehog', *teulada* [taw'laða] 'roof', and *llençol* [ʎan'sɔl] 'sheet' (Carrera i Sabater 2002).

Nonetheless, what is really important is the maintenance of the distinction between *a* and *e*, especially in the final syllable, and its consequences in the inflectional morphology. In Western Catalan formal gender differences in nouns and adjectives based on the endings *-a* and *-e* (*un mestre*, masculine, *una mestra*, feminine, 'teacher') are preserved, whereas in Eastern Catalan they have been lost: Western Catalan *mestr[e]*, masculine, *mestr[a]* (or *mestr[ɛ]*) feminine, compared with Eastern Catalan *mestr[ə]*, masculine and feminine. Furthermore, in Western Catalan the association between the vowel *a* and the feminine gender and between the vowel *e* and the masculine gender is so close that it has brought about linguistic changes. Thus, for example, invariable nominal elements with the ending *-ista* are realised in two ways, as in *ciclista* [si'kliste], masculine, versus [si'klista] (or [si'klistɛ]), feminine, 'cyclist'. We see the same phenomenon with the ending *-aire*, as in *xerraire* [tʃa'rajre], masculine, versus [tʃa'rajra] (or [tʃa'rajrɛ]), feminine, 'talkative'.

2.1.2 The basis for the Western-Eastern divide as proposed by Fabra (1906)

The basis for the Western-Eastern division was reinforced by Fabra (1906, 20–21, fn 2) through the addition of a new criterion based on the evolution of Latin long *E* and short *I* in stressed syllables. While the Classical Latin short *E*, converted into [ɛ] in Vulgar Latin, has generally followed the same evolutionary path in all varieties of Catalan (towards closure in [e], except when followed by *-L-*, *-N'R-*, *-RR-*, *R* + non-labial consonant, and Romance [w] < *-D-*, *-C^e.i*, *-T[j]-*), the Vulgar Latin [e], derived from both long *E* and short *I* of Classical Latin, has followed different paths in the two branches of Catalan. For example, the Catalan proper name *Mercè* (from Latin *MERCĒDE*) is pronounced [mer'se] in Western Catalan but [mər'sə] or [mər'sɛ] in Eastern Catalan. In this dialectal block the transformation of the original [e] into [ə] began around the 13th century, and this pronunciation can still be heard in most of Majorca (the exceptions being the towns of Binissalem, Alaró, Lloseta and, in the process of disappearing, Porreres; Moll 1980, 55–56), in the western half of Minorca, comprising the localities of Ciutadella, Ferreries, Es Migjorn Gran (Moll 1932, 401), in eastern and central Eivissa, including Vila, and in Formentera (Torres i Torres 1983; Veny 1999, 46).

However, outside of these conservative Balearic varieties, this stressed [ə] has proved unstable and has undergone changes in the remainder of the Eastern Catalan domain. In Eastern Catalonia, [ə] had become [ɛ] by the end of Middle Ages, probably for reasons of linguistic economy, given the existence of only a few minimal pairs like *DĒBET* > ['dəw] '(he) owes' and *DĒCEM* > ['dɛw] 'ten' (Veny ¹²1998, 28). The change to

[ɛ] has taken place most recently, during the 19th century, in some Balearic varieties (Moll 1991, 74). In Northern Catalonia, the vowel inventory has shrunk, with the result that all instances of *e*, whether open or closed, have converged in a single mid *e*, [ɛ], which is the phonetic realisation used in that dialect today in words whose etyma contained *Ē* or *Ī*. That said, in the early 20th century both Alcover (1909, 217) and Fouché (1924, XXII) attested to the survival of stressed [ə] in Northern Catalonia in specific localities. Lastly, in l'Alguer, stressed [ə] has returned to [ɛ], probably due to contact with Sardinian, a conservative Romance language which has maintained the [ɛ] it inherited from Vulgar Latin (Kuen 1934, 110–112).

Classical Latin	Vulgar Latin		Preliterary Catalan	Old Eastern Catalan				
Ē, Æ	Ē	+ -L-, -N'R-, -RR-, -R + non-labial cons., [w] < -D-/C ^{e.i} /T[j]-	[ɛ]	[ɛ]	[ɛ]	[ɛ]	[ɛ]	
			MĒL > <i>mel</i> ‘honey’, GĒNĒRU > <i>gendre</i> ‘son-in-law’, TĒRRRA > <i>terra</i> ‘land’, HIBĒRNU > <i>hivern</i> ‘winter’, PĒDE > <i>peu</i> ‘foot’					
		Elsewhere	[e]	[e]	[e]	[ɛ]	[e]	
TĒMPUS > <i>temps</i> ‘time’, PĒTRA > <i>pedra</i> ‘stone’, HĒRBA > <i>herba</i> ‘herb’, PĒLLE > <i>pell</i> ‘skin’, LĒPĒRE > <i>llebre</i> ‘hare’								
Ē, Ī, Æ	Ē		[e]	[ə]	[ɛ]	[ɛ]	[e]	
		CĒPA > <i>ceba</i> ‘onion’, PĪRA > <i>pera</i> ‘pear’, RECĪPĒRE > <i>rebre</i> ‘to receive’, CATĒNA > <i>cadena</i> ‘chain’, PLĒNU > <i>ple</i> ‘full’						
			Western Catalan	Balearic	Centr.	Ross.	Alg.	
				Eastern Catalan				

Figure 2: The evolution of Vulgar Latin vowels *Ē* (< *Ē*, *Æ*) and *Ē* (< *Ē*, *Ī*, *Ē*) in the dialects of Catalan.

It should be noted that in Western Catalan, in certain phonetic contexts propitious to vocalic opening – more or less coinciding with those that have affected the evolution of the short *E* – today we find an open *e* (Hilty 1980). In some cases, the resulting [ɛ] is present in Eastern Catalan too, suggesting that the vocalic opening may have taken place generally and very early (*CĪNĒRE* > *cedra* ‘ash’, with *e* realised as [ɛ] in both Western and Eastern Catalan, including Balearic and Alguerese). However, in other cases the Balearic result with [ə] and the Alguerese result with [e] suggest a restricted (and perhaps late) vocalic opening in Western Catalan (*VĪRĪDE* > *verd* ‘green’, with [ə] in Balearic, [e] in Alguerese and [ɛ] in Central and also Western Catalan). In turn, we occasionally find in Eastern Catalan [e] < *Ē*, *Ī* conditioned by the phonetic context (when the vowel is preceded by [ʒ], when it is followed by [j], [ŋ], [ɲ] or when it is between two palatal consonants as in *LĪGNA* > *llenya* ‘firewood’). In some cases, the closed result is common throughout Eastern Catalan, suggesting an early constraint which could have blocked the evolution into [ə] (*GĪPPU* > *gep* ‘hump’, with [e] in Balearic and in Central Catalan), but in others, the Balearic [ə] points to a late and

restricted closure (DIE DOMÍNĪCU > *diumenge* ‘Sunday’, with [ə] in Balearic, [e] in Algerese and [e] or [ɛ] in Central Catalan). The evolution of these Latin vowels in Catalan is summarised in Figure 2.

2.1.3 The basis for the Western-Eastern divide as proposed by Veny (121998, 1978)

Veny (121998, 19–20) definitively consolidated the Western-Eastern divide by adding a dozen new isoglosses in his best-selling book *Els parlars catalans*, which first appeared in 1978 and still serves as the reference manual for the study of Catalan dialectology. In later works, synthesised in Veny (2015), the author was able to add new lexical isoglosses thanks to the vast amount of material recorded at 190 inquiry points and compiled for the ambitious geolinguistic project *Atles Lingüístic del Domini Català*. The project has already led to the publication of nine volumes of point text maps (Veny/Pons 2001–2018) and six volumes of annotated area maps in colour with three more planned (Veny 2007–2017). Let us now look at the chief discriminating criteria identified by Veny.

First of all, unlike what we see in Eastern Catalan, in Western Catalan the voiceless prepalatal fricative /ʃ/ occurs in neither initial nor post-consonantal position, where we find instead the voiceless prepalatal affricate /tʃ/, as in *xarxa* [ˈtʃartʃa]/[ˈʃartʃə] ‘net’, *xocolata* [tʃokoˈlata]/[ʃukuˈlatə] ‘chocolate’ or *panxa* [ˈpantʃa]/[ˈpaɲʃə] ‘belly’.⁵ Likewise, in Western Catalan, when the voiceless prepalatal fricative /ʃ/ occurs, always in intervocalic or final position, it is preceded by a palatal glide, which is absent in Eastern Catalan: *caixa* [ˈkajʃa]/[ˈkaʃə] ‘box’, *cuixa* [ˈkuiʃa]/[ˈkuʃə] ‘thigh’, *coix* [ˈkojʃ]/[ˈkoʃ] ‘lame’, *peix* [ˈpejʃ]/[ˈpeʃ] ‘fish’.⁶ Next, while in Western Catalan the plural of words derived from Latin proparoxytones ending in -N keep the nasal, in Eastern Catalan the nasal is lost: *HŌMĪNES* > *hòmens/homes* ‘men’, *IŮVĒNES* > *jòvens/joves* ‘youths’, *MARGĪNES* > *màrgens/marges* ‘margins’, *ASĪNOS* > *àsens/ases* ‘donkeys’. Also, the inchoative root extension of some third conjugation verbs (2.1.1 Word Classes, Inflectional Categories and Paradigms) in Western Catalan has a stressed *i*,⁷ in contrast with

5 In Valencian (and less systematically in Northwestern Catalan) we also find the voiced prepalatal affricate /dʒ/ in initial and postconsonantal position, where Eastern Catalan has the voiced prepalatal fricative /ʒ/: *gelat* [dʒeˈlat]/[ʒəˈlat] ‘frozen’, *mengen* [ˈmɛndʒɛn]/[ˈmɛɲʒɛn] ‘(they) eat’.

6 In Northwestern Catalan, which unlike Valencian contains a voiced prepalatal fricative /ʒ/ in its inventory, the palatal glide is also present before it, as in *pluja* [ˈplujʒa] ‘rain’ or the proper name *Roger* [rojˈʒe] ([ˈpluʒə] and [ruˈʒe] respectively in Eastern Catalan).

7 Two origins have been suggested for the Western form. It may have evolved from a particular Latin infix with a long *i*, -ISC- (Badia i Margarit 1981, § 155; Moll 1991, 166) or, alternatively and more probably, it evolved from the general Catalan extension with *e* < -ĒSC-, like similar Western developments in other words with a vowel *e* followed by a voiceless prepalatal sibilant ([ˈejʃ] > [ˈij] and [ejʃ] > [ijʃ]), helped by the glide [j] present in Western Catalan but not in Eastern Catalan. This phenomenon can be seen in the first conjugation verb *deixar* ‘to leave’, derived from *LAXARE*, whose third person

the stressed *e* of Eastern Catalan (< -ESC-), which may be realised phonetically as [ə], [ɛ], [e] or [e] in accordance with the pattern described by Fabra:⁸ *partix* [par'tij] or *parteix* [pər'təj]/[pər'tɛj]/[pər'tɛj]/[pal'tɛj] '(it) splits'. And finally, in Western Catalan the vowels in present subjunctive endings are *a/e* (or, less usual, *o*), while in Eastern Catalan the subjunctive mood has evolved and is now conjugated with *i*:⁹ (*vol que*) *cantes* ['kantes]/(*vol que*) *cantis* ['kantis] '(he/she wants you) to sing', (*vol que*) *perda* ['pərðə]/(*vol que*) *perdi* ['pərði] '(he/she wants me) to lose'.

2.2 The six main dialects of Catalan

At a more specific level than the broad macro-division between Western and Eastern Catalan, the language can be subdivided into dialects and even subdialects and local speech varieties. In 1926, Antoni Maria Alcover proposed a dialectal division of Catalan into six dialects in the introduction to the *Diccionari català-valencià-balear* (DCVB),¹⁰ a comprehensive ten-volume dictionary containing the dialectal and historical Catalan lexicon, with etymologies, documentations and pronunciations, which was promoted and initiated by him but mostly written and ultimately completed by his disciple Francesc de Borja Moll.¹¹ Alcover's proposal served as the basis for what is the currently accepted dialectal division, which is spelled out in Veny (¹²1998, ¹1978).¹²

singular present indicative tense is LAXAT > *lleixa* (Old Catalan) > *deixa* (Modern Catalan, with palatal dissimilation), pronounced ['dijə] (< ['dejja]) in Western Catalan, or the noun *queixal* 'molar' pronounced [ki'jal] (< [kej'jal]) in this same dialectal block (Casanova 1993; Pérez Saldanya 1998, 109–111).

8 With regard to the inchoative root extension, the vowel is not the only difference between dialects. Valencian and Balearic preserve the velar consonant in the first person present indicative tense ending as well as in the subjunctive forms: *servisc*, often pronounced [ser'viʃk], or *serveisc* '(I) serve' as opposed to *servixes* or *serveixes* '(you) serve'. Other dialects have unified the inchoative increment into a single form without a velar consonant: *servixo* '(I) serve' like *servixes* '(you) serve' or *serveixo/serveixi/serveix* '(I) serve' like *serveixes* '(you) serve'.

9 See Gulsoy (1993, ¹1976) and Pérez Saldanya (1998, 147–168) for the history of the development of the variants of the present subjunctive endings in Catalan.

10 The introduction contained in the volume 1 (1930) of the DCVB, first appeared in 1926 in a fascicle, is signed by Alcover and Moll. However, it includes, with some tweaks, the dialectal division of Catalan previously proposed by Alcover (1902–1903, 1918–1919). In the second corrected and improved edition of the first volume of the DCVB, dated in 1968, the abovementioned introduction was replaced by another, written by Moll, which lacks the dialectal division of Catalan.

11 This dictionary was complemented by the publication between 1980 and 1991 of the nine-volume *Diccionari etimològic i complementari de la llengua catalana*, a compendium of Catalan dialectal and historical data compiled by Joan Coromines with the assistance of Joseph Gulsoy and Max Cahner. Before his death in 1997, Coromines managed to write a final supplementary volume, which was published in 2001 along with an index of forms for the full work prepared by Joan Ferrer, Josep Ferrer and Joan Pujadas.

12 However, the labels proposed by Alcover for the dialects (and used in the DCBV) do not always coincide with the current nomenclature. Thus, what he labels "Eastern Catalan" is now called *Central*

2.2.1 The dialects of Catalan as proposed by Alcover (1926)

The current dialectal division of Catalan derives from Alcover's observation that the first person present indicative tense ending for first conjugation verbs – a very frequent form in the spoken language – exhibits five variants across the Catalan linguistic domain.¹³ The historical origin of these variants is one of the most thoroughly studied topics in Catalan linguistics (Par 1930; Coromines 1971; Blasco Ferrer 1985; Casanova 1989; Gulsoy 1993, 1987; Alsina 1988–1989; Wheeler 1995; Pérez Saldanya 1998, 127–146).

The oldest of these five forms is preserved in Balearic and Alguerese, two dialects isolated by geography (and history in the latter case): CANTO > *cant* '(I) sing', with a regular apocope, which in Preliterary Catalan affected final unstressed vowels other than -A (Rasico 1982, 114). The evolutionary next step was final -e, as in *cante* '(I) sing', which can still be found in Contemporary Valencian. The presence of this vowel has been explained as a morphologisation from an epithetic vowel at first only added to verbal forms which violate the sonority scale or other syllabic formation rules as in *sembl* '(I) look like', *sembre* '(I) sow', *entre* '(I) go in', *parle* '(I) speak', *lliure* '(I) deliver' and *erre* '(I) miss'.¹⁴ In continental Catalan this vowel was slowly extrapolated to all of the other first conjugation verbs, starting in Valencia in the 15th century and spreading outwards (Pérez Saldanya 1998, 132–134).¹⁵ These verbal forms ending in -e fit well within the Valencian verbal paradigm, since they serve to differentiate first from third person present indicative tense endings for first conjugation verbs, as in *cante* '(I) sing' versus *canta* '(he/she) sings'. The same cannot be said, however, about either Northern

Catalan, "Western Catalan" is now *Northwestern Catalan* and "Eastern-Pyrenean Catalan" is now *Northern Catalan*.

13 In fact, used alone, this criterion yields only five Catalan dialects (Balearic, Valencian, Northern Catalan, Northwestern Catalan and Central Catalan), because in Alguerese the first person present indicative tense ending is identical to that used in Balearic. However, because Alguerese is the dialect of a single community geographically, historically and politically isolated from the rest of the Catalan linguistic domain, it has conserved numerous linguistic traits that have become archaic in other dialects and, for the same reason, Alguerese has produced several linguistic innovations unknown in the other Catalan-speaking areas.

14 It should be noted that historically this epithetic vowel was also present in these verbs in Balearic and Alguerese, but over time it was lost to avoid allomorphy between forms with and without vowel (Alsina 1988–1989, 102, 104–107): *sembl* '(I) look like', *compr* '(I) buy', *parl* '(I) speak'. As a result, these dialects currently contrast between verbal forms that lack the support vowel and other words where the vowel is present: *entr* '(I) go in' but *entre* 'between' (preposition), *m'alegr* 'I am glad' but *alegre* 'happy' (adjective) (Pérez Saldanya 1998, 135).

15 This development did not extend to second and third conjugation verbs; even now, like Balearic and Alguerese, Valencian has no vocalic marker in these conjugations. However, unlike the other two dialects, Valencian velarises the verbal root in first person forms: from *perdre* 'to lose', *perd* '(he/she) loses' but *per[k]* '(I) lose', from *sentir* 'to hear', *sent* '(he/she) hears' but *sen[k]* '(I) hear' (Wheeler 1984). Valencian is the only contemporary dialect in which first person present indicative tense first conjugation forms differ from those of the second and third conjugations.

and Central Catalan on the one hand, because by the 13th century unstressed *a* and *e* had already converged in [ə] in these two dialects, or about Northwestern Catalan on the other, given that this dialect has [e] as a third person present tense ending for first conjugation verbs (*cant[e]* ‘(he/she) sings’).¹⁶ It is not surprising, then, that in these dialects subsequent developments have occurred.

Nowadays Northwestern and Central Catalan have the same phonological vowel marker for the first person present indicative tense ending, /o/, but it is realised as [o] or [u] respectively, as noted above. This -o, documented in some verbs as far back as the 15th century (Gulsoy 1993, 1987), may well have started out as assimilation – a common phonetic process with many examples in Catalan – in verbs with a stressed velar vowel (Coromines 1971) as in *compre* > *compro* ‘(I) buy’, *cobre* > *cobro* ‘(I) cash’, *mostre* > *mostro* ‘(I) show’, *compte* > *compto* ‘(I) count’ and *dubte* > *dubto* ‘(I) hesitate’. Later, speakers would have taken advantage of this phonetic tendency to prevent syncretism by restricting the [o]/[u] ending to the first person and reserving the [e]/[ə] ending for the third person.¹⁷ Once this change was well established in first conjugation verbs, the ending -o was extended to the other conjugations, in a process that began in the 17th century and came to a close in the 19th (Gulsoy 1993, 1987).

Northern Catalan has developed another solution. Today, the first person present tense ending for all conjugations is -i, as in neighbouring Occitan. This -i may initially have reflected a preservation of the yod from the -iō ending of Latin fourth conjugation verbs in cases in which it facilitated the pronunciation of consonant sequences not allowed in word-final position (Coromines 1971): *DORMIō* > *dorm* ‘(I) sleep’, *SENTIō* > *sent* ‘(I) hear’ but **OPRIō* > *obri* ‘(I) open’, **SOFFRIō* > *sofri* ‘(I) suffer’. Maintaining it would have been favoured both because of its coincidence with the thematic vowel of this conjugation and also because of its presence in the old present subjunctive forms ending in -ia like *òbria* or *sòfria* (Gulsoy 1993, 1987). Later this

16 Two hypotheses have been formulated about the origin of this ending [e] derived from Latin -AT. Being one of the most characteristic traits of Northwestern Catalan, it is found not only in first conjugation present indicative (*cant[e]* ‘(he/she) sings’) but also in the imperfect indicative (*cantav[e]* ‘(he/she) sang’) and conditional (*cantari[e]* ‘(he/she) would sing’) of all three conjugations, and in second and third conjugation present subjunctive (*volen que perd[e]* ‘they want him/her to lose’). The first hypothesis is phonological and postulates a late loss of the final consonant, which led to the vocal closing -AT > -[et] > -[e] in Northwestern Catalan, as opposed to -A(T) > -[a] in all other Catalan dialects (Rasico 1982, 95, fn 276; Wheeler 2007, 108). The second hypothesis is morphological and suggests, on the basis of ancient documents, a common origin with final [a], which was later replaced by [e] in Northwestern Catalan – the dialect most reluctant to syncretism – to avoid conflation between the first and third person in imperfect indicative (1*l**cantava*, 2*l**cantaves*, 3*l**cantava* → 3*l**cantave*, 6*l**cantaven*), and later this change was extended to other tenses like present indicative (1*Pl**cant*, 2*Pl**cantes*, 3*Pl**canta* → 3*Pl**cante*, 6*Pl**canten*) because of pressure for congruity in the verbal paradigm (Massanell i Messalles 2016).

17 This morphological change could have been encouraged by the fact that this -o coincides with the ending found in other Romance languages like Castilian and Italian, and also with the Latin ending (Par 1930; Blasco Ferrer 1985).

vowel ending spreads first to the verbs of the same conjugation that do not need it for phonetic reasons like *dormi* ‘(I) sleep’ or *senti* ‘(I) hear’, and later to the first conjugation, replacing the *-e* and disambiguating the syncretism between first and third person: *canti* ‘(I) sing’ but *cant[ə]* ‘(he/she) sings’.

Several observations are necessary here. First of all, the *-o* ending has managed to spread across a wide area of the North of the Valencian Country, which is consistent with the fact that this area, together with Southern Catalonia, is a major dialectal transition area crossed by the various isoglosses that mark the gradual transition from Northwestern Catalan to Valencian. Alcover himself – not for linguistic reasons but rather to avoid having to sever the North of the Valencian Country from the Valencian dialect region – suggested that, in this case, a further linguistic criterion was needed to establish the border between Northwestern Catalan and Valencian: the imperfect subjunctive marker. This isogloss, which closely approximates the administrative boundary between the communities of Catalonia and Valencia, separates the northern sigmatic forms like (*volia que*) *cantés* ((he/she) wanted me) ‘to sing’), *perdés* (‘to lose’), and *sentís* (‘to hear’), from the analogous southern rhotic forms *cantara*, *perdera* and *sentira*, and it has become the standard dividing line between these two Catalan dialects.

Secondly, when Alcover collected data on Catalan verbal inflections at the beginning of the 20th century, ultimately creating a compendium that is even today regarded as an exceptional example of its kind in any of the Romance languages (Alcover/Moll 1929–1932), the ending *-i* was found both in Northern Catalonia, under French administration, and also in a strip south of the administrative border including the regions of Alt Empordà, Garrotxa, Ripollès and Baixa Cerdanya, under Spanish administration, which was considered as belonging dialectally to Northern Catalan. Now, a century later, because of the official standardisation process that has been undertaken in (Southern) Catalonia – but not in Northern Catalonia – the isogloss has shifted northward to more closely coincide with the administrative border (Adam Aulines 2006, 87–101).¹⁸

Thirdly, there exists another first person present indicative tense ending, with the vowel [a], which is used by a tiny minority of Catalan-speakers living in a small area of the north of the Valencian Country.¹⁹ Here, the contrast between first and third person is *cant[a]* ‘(I) sing’ versus *cant[e]* ‘(he/she) sings’ (Oliver Barreda 2014).

Finally, in some local varieties a velar consonant is added after the vowel, yielding *cante[k]*, *canti[k]* or *cant[uk]* ‘(I) sing’. And in some northern regions of Central Catalan, especially those from which *-i* has disappeared, it is usual to reinforce the vowel [u] with a dental consonant, as in *cant[ut]* ‘(I) sing’.

¹⁸ For this reason, in the current dialect map of Catalan the linguistic variety spoken in this strip is still labelled *Transitional Northern Catalan*, though it is included within the domain of Central Catalan.

¹⁹ This phenomenon affects Figueroles, Costur, les Useres and l’Alcora, all towns in the region of Alcatén.

2.2.2 The dialects of Catalan as proposed by Veny (¹²1998, ¹1978)

Since Alcover, although various hypotheses for the dialectal divisions of Catalan have been formulated (e.g. Barnils 1919; Badia i Margarit 1951, § 20–22; Coromines 1954, 56–57; *Gran Enciclopèdia Catalana* 1973, s. v. *català*; Sardà/Guiter 1975), Veny's proposal (¹²1998, ¹1978) is the one that is currently considered definitive. This eclectic classification situates each variety in one of the two main blocks (Western or Eastern Catalan) and then, within that block, in one of six dialects (Northwestern Catalan, Valencian, Northern Catalan or Rossellonese, Central Catalan, Balearic or Alguerese), including subdialects from transition areas like Transitional Northern Catalan and Tortosan. It also notes relatively localised varieties bound by a single isogloss, such as what is known as *Xipella*, a group of Central Catalan varieties in which -e has evolved into [i] in final unstressed syllables, as in *quinze* ['kinzi] 'fifteen', *vaques* ['bakis] 'cows' and *parlaves* [pər'laβis] '(you) spoke'. It also includes local varieties individualised by a particular historical circumstance, like the Majorcan varieties spoken in the Valencian regions of Marina Alta and Marina Baixa due to resettlement activity in the 17th century (Beltran Calvo 2011; Beltran Calvo/Herrero Lloret 2011a; 2011b). In Figure 3, we reproduce a table showing the dialectal divisions of Catalan from Veny/Massanell i Messalles (2015).²⁰

Blocks	Dialects	Subdialects and localised varieties
Eastern Catalan	Northern Catalan or Rossellonese	Capcinese
		Transitional Northern Catalan
		<i>Salat</i>
	Central Catalan	<i>Xipella</i>
		Barcelonese
		Tarragonese
	Balearic	Majorcan
		Minorcan
		Eivissenc-Formenteran
	Alguerese	

²⁰ The only differences with respect to Veny (¹²1998, ¹1978) are the inclusion of Alacantine as a variety within Southern Valencian and the preference for the name *Central Valencian* rather than the traditional *Apitxat*. This term (*apitxat* < (*a*)*pitjat* 'pressed') refers to the devoicing of sibilants characteristic of this subdialect and has a pejorative connotation, as noted by Colomina i Castanyer (1999, 71).

Blocks	Dialects	Subdialects and localised varieties
Western Catalan	Northwestern Catalan	Ribagorçan
		Pallarese
		Tortosan
	Valencian	Northern Valencian
		Central Valencian or <i>Apitxat</i>
		Alacantine
		Southern Valencian
		Majorcan of Tàrbena and la Vall de Gallinera

Figure 3: The dialectal divisions of Catalan (from Veny/Massanell i Messalles 2015)

3 Description of the Catalan dialects

Due to space constraints, we will focus in the following sections on the characteristics of the six main dialects of Catalan, and only occasionally refer to subdialects or local varieties.

3.1 Western Catalan dialects: Northwestern and Valencian

As we have seen, Northwestern Catalan and Valencian share some crucial traits which justify their inclusion in the same dialectal block, exemplified by *par[e]* ‘father’, *c[o]rder* ‘lamb’, *c[e]ba* ‘onion’, *[tʃ]ic* ‘little’, *co[j]f* ‘lame’, *hòmens* ‘men’, *partix* ‘(it) splits’, *(vol que) perda* ‘(he/she wants) me to lose’. However, the two dialects also exhibit differences between them, such as the previously noted imperfect subjunctive marker, as in *(volia que) cantés/cantara* ‘(he/she wanted) me to sing’, and first and third person singular endings, as in *canto/cante* ‘(I) sing’, *perdo/per[k]* ‘(I) lose’, *cante/canta* ‘(he/she) sings’, *cantav[e]/cantava* ‘(he/she) sang’, *cantari[e]/cantaria* ‘(he/she) would sing’. There are other differences, which we will now detail. As we do so, it is important to bear in mind that the isoglosses corresponding to each of these linguistic variables bisect the map along different lines across a very wide strip of linguistic transition including Tortosan – the southernmost northwestern subdialect – and Northern Valencian (Gimeno Betí 1997).²¹

21 Our knowledge of Tortosan Catalan has been greatly enriched by the recently published regional linguistic micro-atlas produced by Pere Navarro and colleagues at the Universitat Rovira i Virgili (Navarro Gómez 2017).

A good number of the Valencian varieties retain /v/ in their phonemic inventory, while in Northwestern Catalan this phoneme has been conflated with /b/ (realised as [β] in intervocalic position), as in *vaques* ['vakes]/['bakes] 'cows' or *cavall* [ka'vaʎ]/[ka'βaʎ] 'horse'. Northwestern Catalan has eight sibilants:²² /s/, *caçar* [ka'sa] 'to hunt'; /z/, *casa* ['kaza] 'house'; /ts/, *plats* ['plats] 'dishes'; /dz/, *dotze* ['dodze] 'twelve'; /ʃ/, *faixa* ['fajʃa] 'corset'; /ʒ/, *pluja* ['plujʒa] 'rain'; /tʃ/, *cotxe* ['kotʃe] 'car'; and /dʒ/, *metge* ['medʒe] 'doctor'. By contrast, Valencian has neither /ts/, which has been palatalised (*tots* ['totʃ] 'everybody'), nor /ʒ/, because the historical process of deaffrication of /dʒ/ in some contexts has not occurred in this dialect (*roja* ['rɔdʒa] 'red'). Moreover, there are two subdialects, Ribagorçan (within Northwestern Catalan) and Central Valencian that have devoiced all voiced sibilants yielding, for example, *rosa* ['rɔsa] 'rose' and *germà* [tʃer'ma] 'brother'.

Most of the Valencian varieties maintain final *r*, while Northwestern Catalan loses it in all infinitives and polysyllabic nouns: *fer* ['fer]/['fe] 'to make', *parlar* [par'lar]/[par'la] 'to speak', *fuster* [fus'ter]/[fus'te] 'cabinetmaker', *carrer* [ka'rer]/[ka're] 'street'. In Valencian the final consonantal groups of nasal or liquid plus occlusive do not undergo simplification, while in Northwestern Catalan they do: *camp* ['kamp]/['kam] 'field', *banc* ['banʃ]/['ban] 'bank', *pont* ['pɔnt]/[pɔn] 'bridge', *molt* ['molt]/[mol] 'much'. The intervocalic [z] of the ending *-esa* (< -ĪTĪA) and [δ] of the endings *-ada* (< -ATA) and, less systematically, *-ador* (< -ATORE) are dropped in Valencian but are maintained in Northwestern Catalan: *pobresa* [po'βrea]/[po'βreza] 'poverty', *arracada* [ara'ka]/[ara'kaða] 'earring', *bufador* [bufa'or]/[bufa'δo] 'fan to blow the fire'.

Northwestern Catalan retains the classical form of the masculine definite article (whose use has been extended to proper nouns) and of the third person masculine weak pronoun, both coming from the Latin demonstrative ĪLLU, ĪLLOS > *lo*, *los*, while Valencian uses *el*, *els*, created from the classical forms by phonosyntax in vocalic contexts (*lo pare* 'the father' but *la mare e-Ī pare* 'the mother and the father' → *el pare* 'the father'):²³ *lo llibre/el llibre* 'the book', *los peus/els peus* 'the feet', *lo Joan/Joan* (without article before the proper noun in Valencian), *lo busquen/el busquen* '(they) search for him', *lo deixaré allí/el deixaré allí* '(I) will leave it there'. In Valencian the demonstrative paradigm retains three degrees and forms for first and second person that come directly from the Latin demonstratives ĪSTE and ĪPSE: *este* 'this (near me)', *eixe* 'that (near you)'. By contrast, Northwestern Catalan has a two-degree demonstrative paradigm and a single proximity form embracing both first and second person which comes from the reinforced Latin demonstrative *ACCU-ĪSTE: *aquest* 'this (near

²² In fact, though /ts/ has no phonemic value (Badia i Margarit 1988, § 267–271), it has traditionally been included in the inventory of Catalan sibilants. We find this realisation in the plurals of nominal elements ending in *-t* (*gots* '(drinking) glasses', *malvats* 'wicked', *tots* 'everybody'), in the second person singular present indicative tense ending of the verb *poder* (*pots* 'you (sing.) can') and in the adverb *potser* 'maybe' (formed through grammaticalisation of the verb phrase *pot ser* 'it may be').

²³ In Old Catalan the copulative conjunction was *e* (< ET), which later evolved into *i*.

me); that (near you)'. In both systems the distal form is *aquell* (< *ACCU-İLLE) 'that (near him/her/them)'.

Valencian has inverted the medieval order of pronominal combinations of a third person accusative and a third person dative (DO + IO → IO + DO): *done [els llibres]_{DO} [a Daniel]_{IO}* '(I) give the books to Daniel' → *[li]_{IO} [ls]_{DO} done* '(I) give them to him'; Northwestern Catalan, like all other Catalan dialects, has maintained the order (DO + IO) but the dative pronoun has been reduced to [i] by lateral dissimilation: *done [els llibres]_{DO} [al Daniel]_{IO}* '(I) give the books to Daniel' → *[los]_{DO} [hi]_{IO} dono* '(I) give them to him' (Vallcorba i Rocosa 2011).²⁴ In Northwestern Catalan, unlike Valencian (and all other Catalan dialects), unaccusative constructions do not exhibit number agreement between the verb and the indefinite nominal phrase that follows it: *Ve pluges/Venen pluges* 'There come rains', *Enguany ha arribat molts turistes/Enguany han arribat molts turistes* 'This year many tourists have arrived', *Es ven peres/Es venen peres* 'Pears are sold' (Solà 1987; Rigau 1992; 1997).

3.2 Eastern Catalan dialects: Northern, Central, Balearic and Alguerese

In contrast to the two dialects making up Western Catalan, Eastern Catalan contains four: Northern, Central, Balearic and Alguerese. We have seen the common features which distinguish these dialects from Western Catalan, exemplified by *par[ə]/par[a]* 'father', *c[ə]ba/c[ɛ]ba/c[ɛ̃]ba/c[e]ba* 'onion', *[ʃ]afar* 'to flatten', *co[ʃ]* 'lame', *joves* 'young men or women', *parteix* '(it) splits', *(vol que) perdi* '(he/she wants) me to lose'. Naturally, these four dialects also differ among themselves, as we have seen with the variations in the first person present tense ending like *cant/cant[u]/canti* '(I) sing'. Let us now summarise some of the more important features that distinguish these dialects from each other.

Regarding Balearic,²⁵ though there are noteworthy differences between the varieties spoken in Majorca, Minorca and the Pityusic Islands, as we have already seen with regard to stressed vocalism (with or without /ə/) and unstressed vocalism (with or without [o]), we will focus on what these varieties have in common, such as the absence of a vowel in first person present indicative tense endings as noted above.

²⁴ When this lateral dissimilation occurs ([lozli] > [lozi]), the dative pronoun is graphically represented by *hi*, coinciding with the spelling of the locative pronoun *hi* < HIC: *porto les claus a la Núria* '(I) bring the keys to Núria' → *les hi porto* '(I) bring them to her' like *hi anem* '(we) come there'.

²⁵ There exists an audio-visual corpus of Balearic (*Arxiu audiovisual dels dialectes catalans de les Illes Balears*), produced under the direction of Jaume Corbera at the Universitat de les Illes Balears, which contains one hundred hours of video recording. A DVD anthology based on this corpus and published in 2003 consists of 67 excerpts from interviews – totalling around eleven hours of footage – with accompanying phonetic and orthographical transcriptions.

First of all, the linguistic conservatism of Balearic is exemplified in its consonant system. Thus, on the one hand, the voiced labiodental fricative /v/, coming from Latin V-, -V-, -B- and -PH-, has not been conflated with /b/ as it has in other Catalan dialects, but rather has been maintained: VĪNU > vi ['vi] 'wine', CAVARE > *cavar* [kə'va] 'to dig', FABA > *fava* ['favə] 'bean', CŌPHĪNU > *cove* ['kɔvo]/['kɔvu]/['kɔvə] 'basket'.²⁶ And, on the other hand, the Latin groups -L[j]-, -C'L- (sometimes attracting -T'L- and -P'L-) and -G'L- in Balearic yield the result [j], and not [ʎ] like in most areas of the Catalan linguistic domain, where the evolution of these groups coincides with the result of L- and -LL-: FŌLĪA > *fulla* ['fujə] 'leaf', ŌC(Ū)LU > *ull* ['uj] 'eye', RĒG(Ū)LA > *rella* ['rəjə] 'ploughshare', VĒT(Ū)LA → VEC'LA > *vella* ['vejə] 'old woman', RESTŪP(Ū)LU → REST-UC'LU > *restoll* > *rostoll* [ros'toj]/[rus'toj] 'stubble'. In final position when following an i, [j] is absorbed, CŪNIC(Ū)LU > *conill* [ko'ni]/[ku'ni] 'rabbit', and it can also disappear in intervocalic position, ['vejə] > ['veə], following a palatal vowel, and even ['fujə] > ['fuə], following a velar vowel.

With regard to morphology, a striking feature is the preservation of the personal articles *en*, *na*, *n'* (as in *en Joan*, *n'Antoni*; *na Margalida*, *n'Aina*), which come from the Latin vocatives DOMINE 'lord' and DOMINA 'lady'. So too have survived with full vitality the definite articles coming from the Latin intensifier demonstratives ĪPSU/ĪPSA/ĪPSOS/ĪPSAS '(him)self/(her)self/(them)selves', *es*, *sa*, *s'*, *ses* (as in *es matí* 'the morning', *s'estel* 'the star', *sa llebre* 'the hare', *s'hora* 'the hour', *es cans* 'the dogs', *ses eines* 'the tools'), with the single exception of Pollença, in Majorca, where the article coming from ĪLLU/ĪLLA/ĪLLOS/ĪLLAS is used. The *es/sa/s'/ses* definite articles show variation in the different islands in accordance with particular phonetic tendencies. Thus, in both Majorcan and Minorcan, forms ending in -s undergo affrication when they precede another sibilant (Pons Moll 2007), as seen in *es joves* [əd'dʒovəs] 'the youths (young people)' and *ses sabates* [sətsə'βatəs] 'the shoes'; the expansion of this affricate pronunciation is likely the origin of the variant *ets* used in these two subdialects before a masculine plural noun beginning with a vowel, as in *es homes* [əd'dʒɔmos]/[əd'dʒɔmus] 'the men' as opposed to Eivissenc-Formenteran [ə'zɔməns].²⁷

All these Balearic phenomena also persist in particular locations within the Central Catalan domain, but in a clearly recessive way. For example, within the Tarragonese subdialect, the phoneme /v/ is maintained in the regions of Alt Camp and Baix Camp and in the periphery of the city of Tarragona, though no longer in the

²⁶ The three transcriptions correspond respectively to Majorcan, Minorcan and Eivissenc-Formenteran. The forms with a final -o (pronounced [o] or [u] as described above) are due to velar assimilation to the preceding stressed vowel.

²⁷ The survival of the nasal in plurals derived from Latin proparoxytones in this Eastern Balearic subdialect is due to the frequent contact between the populations of Eivissa and the Valencian coast due to maritime commerce. This nasal is also preserved in Tarragonese, an Eastern Central subdialect with transitional traits from Northwestern Catalan. These exceptions do not invalidate the general claim that the nasal is usually found in Western Catalan but not in Eastern Catalan.

city itself. Noteworthy here is the fact that /v/ can also come from a Latin -P- by bilabial dissimilation, as in PUPĪLLA > *pubilla* [pu'viɫə] 'heiress' and, with a preceding metathesis, PAUPĒRE > *pobre* ['prɔvə] 'poor' and PIPĒRĪNA > *pebrina* [prə'vinə] 'pepper'. By the same token, the [j] resulting from -L[j]-, -C'L- or -G'L- can still be found in some Central Catalan varieties, especially in rural areas and among elderly speakers, though it is clearly recessive and is stigmatised as being quaintly old-fashioned. The geographic distribution of this phonetic realisation varies depending on the word: CĪLĪA > *cella* ['sejə] 'eyebrow' is very widespread and extends even to the capital city Barcelona, while COCHLEARIA > *cullera* [ku'jerə] 'spoon' and ACŪCŪLA > *agulla* [ə'ɣujə] 'needle' have a more restricted range (Veny 1984, 45–46; 2007–2017, vol. 1, maps 19 and 20). In Central Catalan the personal article *en* survives before a personal noun beginning with a consonant, but the rest of the paradigm has been lost: *en Jordi* (as well as *el Jordi*), but only *l'Antoni*, *la Núria*, *l'Eulàlia*. The so-called *Salat* ('Salty') variety of Central Catalan, which includes the local speech of Cadaqués in the region of Alt Empordà and the varieties spoken in other coastal villages between the rivers Ter and Tordera, is a subdialect isolated within Central Catalan which preserves the definite articles *es/sa/s'/ses* 'the' derived from the Latin intensifier demonstratives ĪPSU/ĪPSA/ĪPSOS/ĪPSAS '(him)self/(her)self/(them)selves' (and otherwise limited to Balearic, as noted above). However, the inhabitants of these villages use it only among themselves, rarely with outsiders. Nonetheless, a few place names here and there throughout the Central Catalan domain attest to the former vitality of this definite article, which appears agglutinated in *Sant Climent Sescebes* (*ses cebes* 'the onions'), *Sant Joan Despi* (*des pi* 'from the pine'), *Sant Just Desvern* (*des vern* 'from the alder') and *Collserola* (*s'erola* 'the little threshing-floor').

With regard to Northern Catalan (Gómez Duran 2016), a dialect strongly influenced by Occitan adstratum – especially in the region of Capcir – and French superstratum, one striking feature is that large areas have reduced the tonic vowel system from eight elements to five. The first step in this process was the closing of /o/ derived from Latin long O and short U to /u/ (Costa 1977; Escudero 1999) as seen in the toponyms *Canigó* [kəni'ɣu], *Rosselló* [rusi'ɫu], and in *mosca* ['muskə] 'fly', *llop* ['ɫup] 'wolf'. This linguistic change, which was already underway in Occitan in the 14th century, is documented in Northern Catalan in the 17th century through spellings like *mussa* (1627, *mossa* 'teenage girl'), *gus* (1639, *gos* 'dog') and, by hypercorrection, *arcabossos* (1613, *arcabussos* 'arquebuses'), *llegoms* (1641, *llegums* 'legumes') (Veny 1980, 473–474). Once the /o/ > /u/ closure had been consolidated, the remaining transformations came about to balance the system: open *o* became a mid *o* equidistant between /a/ and /u/ (*dona* ['dɔnə] 'woman'), and later open *e* and closed *e* became a mid *e* (*cel* ['sɛl] 'sky', *llet* ['ɫɛt] 'milk'), thus restoring the same degrees of aperture in front vowels as in back vowels.

There are several phonetic realisations that are particularly characteristic of Northern Catalan. Examples include the loss of final unstressed vowels in proparoxytones ending in *-ia* (*família* [fə'mili] 'family', also current in some Balearic varieties)

and the reduction of the voiceless prepalatal fricative /ʃ/ to [j], but only when it occurs in final position (*mateix* [mə'tɛj] 'same', masculine singular, but *mateixa* [mə'tɛʃə] 'same', feminine singular). Noteworthy in terms of nominal morphology is the loss of the nasal in plurals derived from Latin paroxytones ending in -N' (PANES > *pans* > *pas* 'breads', CAMĪNOS > *camins* > *camís* 'ways'). So too is the survival – perhaps favoured by contact with French – of the third person plural possessive adjective *llur* ['ʎurt] (pronounced with a dental reinforcement like all monosyllables with final *r*), *llurs* ['ʎurs], and also *llura* ['ʎurə] and *llures* ['ʎurəs], the feminine forms created by analogy. Thus, while *el seu hort* means 'his/her vegetable garden', *el llur hort* means 'their vegetable garden'. All the other Catalan dialects have long since replaced these forms with the third person singular possessive adjectives *seu/seus/seva* or *seua/seves* or *seues*.²⁸

Regarding verbal morphology, in Northern Catalan the vowel -i of the first person present indicative tense ending (*canti* '(I) sing') has been extended to the first person forms of other verbal tenses, thus replacing the historical endings with final unstressed -a: CANTABAM > *cantava* → *cantavi* '(I) sang', CANTARE HABEBAM > *cantaria* → *cantariï* '(I) would sing'. The ending -i is present in all Eastern Catalan varieties in the present subjunctive tense, as in (*vol que*) *canti* '(he wants) me to sing',²⁹ but Northern Catalan has extended this -i to the imperfect subjunctive tense as well, as in *cantés* → *cantessi*, (*volia que*) *cantessi* '(he wanted) me to sing'. With regard to syntax, the formation of negative sentences in Northern Catalan warrants comment. The negative particle in Catalan is *no* 'not' (e.g., *No vindré* 'I will not come'), and throughout Catalonia (but not in the Valencian Country and the Balearic Islands) this particle can be emphatically reinforced by *pas* when the speaker contradicts an assumption (*Després del que ha passat, no vindré pas!* 'After what has happened, I (certainly) will not come!'). In Northern Catalan – again perhaps due to the French influence – the use of *pas* has been extended to all negative sentences, emphatic or not, and this has facilitated the loss of *no*, this particle having become redundant: *Vindré pas* 'I will not come', *Ho toquis pas!* 'Do not touch it!', *Ho sem pas trobat* 'We have not found it', *Tenes pas sal?* 'Don't you have any salt?' (Gómez Duran 2016, 140–150).

In Alguerese (Scala 2003), due to its isolation, two opposing tendencies converge: the preservation of linguistic variants that have become archaic in other dialects, and the appearance of linguistic innovations unknown in the rest of the Catalan linguistic domain. Regarding preserved elements, Alguerese shares with the other consecutive dialects (Valencian and Balearic): the presence of /v/ in the consonant inventory, as in *vi* ['vi] 'wine'; the unreduced pronunciation of final groups of nasal or liquid plus

²⁸ *Llur* and *llurs* are preserved in Standard Catalan, especially in formal written contexts. In Alguerese these possessive forms have ultimately been replaced with the prepositional phrases *d'ellos*, masculine, and *d'elles*, feminine, literally 'from them'.

²⁹ Forms like *càntia* were presumably generated by analogy with historical forms ending in -ia like *sàpia* (< SĀPIAM) and *càpia* (< CĀPIAM), to later become *canti* by apocope (Gulsoy 1993, 1976).

occlusive, as in *pont* [ˈpɔnt] ‘bridge’, *alt* [ˈalt] ‘high’; and the absence of palatalisation of *-tl-* from *-T’L-*, *-D’L-* and *-[j]’L-*, like SPATŪLA > *espatla* [asˈpala] ‘shoulder’ and Vulgar Latin AMIDD(U)LA > *ametla* [ˈmela] ‘almond’. We can also include in this conservative tendency the maintenance of etymological forms of certain grammatical particles: the masculine definite articles ILLU > *lo* [lu], ILLOS > *los* [lus] ‘the’, as in Northwestern Catalan; the clitic pronouns ME > *me* [ma] ‘me’, TE > *te* [ta] ‘you (sing.)’, SE > *se* [sa] ‘himself/herself/themselves’, INDE > *ne* [na] ‘it’, NOS > *nos* → *mos* [mus] ‘us’, VOS > *vos* [vus] ‘you (pl.)’, as in Western Catalan and some minority areas of Eastern Catalan;³⁰ and the feminine forms of singular possessives MĚAM > *mia* ‘my, mine’, TŪAM > *tua* ‘your, yours’, SŪAM > *sua* ‘his, her, hers’, replaced in all the other dialects with forms created from masculine: MĚUM > *meu* and, then, *mia* → *meua*, this last form being transformed into *meva* [ˈmeβə] in Central Catalan (and [ˈmevə] in some Balearic varieties) on the basis of pairs like *blau/blava* ‘blue’, *tou/tova* ‘soft’.

With regard to linguistic innovations in Alguerese, various phonetic transformations seem to have been favoured due to contact with Sardinian. Beyond the previously mentioned loss of the shwa, which has become /e/ in stressed syllables and [a] in unstressed ones, as in PĪRA > *pera* [ˈpərə] > [ˈpera] ‘pear’ and RECĪPĒRE > *rebre* [ˈrəβrə] > [ˈrebra] ‘to receive’, several consonantal changes have also occurred. Thus, intervocalic *d* and *l* have become [r], as in *gelada* [dʒaˈrara] ‘frost’, *vida* [ˈvira] ‘life’, *vila* [ˈvira] ‘town’; *l* as the second element of a consonant group opening syllable has become [r], as in *flassada* [fraˈsara] ‘blanket’, *blau* [ˈbraw] ‘blue’; implosive *r* has become [l], as in *tarda* [ˈtalda] ‘afternoon’, *març* [ˈmals] ‘March’, *obert* [uˈbɛlt] ‘open’, and this change also affects words ending in *-rm*, which is reduced to the lateral, as in *forn* [ˈfol] ‘oven’, *infern* [imˈfɛl] ‘hell’, as well as the plural of words ending in *-r*, as in *flor* [ˈfrɔr] ‘flower’ but *flors* [ˈfrɔls] ‘flowers’, even when this consonant has become silent in the singular, as in *servidor* [salviˈro] ‘servant’ but *servidors* [salviˈrols] ‘servants’. In addition, the palatal lateral /ɬ/ and palatal nasal /ɲ/ lose their palatal character when they are in word-final position: *fill* [ˈfil] ‘son’, *fills* [ˈfils] ‘sons’ but *filla* [ˈfiɬa] ‘daughter’, *filles* [ˈfiɬas] ‘daughters’, and *any* [ˈan] ‘year’, *anys* [ˈans] ‘years’ but *anyada* [aˈɲara] ‘vintage’.

4 Concluding remarks

To summarise, there are various fundamental linguistic variables at play in the dialectal diversity of Catalan.³¹ Variables at the phonetic and phonological levels include: the inventory of stressed vowels, with five elements (i/e/a/o/u), seven (i/e/ɛ/a/ɔ/o/u) or

³⁰ Instead of these, in Central Catalan there exist more evolved forms that have resulted from phonosyntax in vocalic contexts (*te busca* ‘he looks for you’ but *diu que·t busca* ‘he says he looks for you’ → *et busca* ‘he looks for you’): *em* ‘me’, *et* ‘you (sing.)’, *es* ‘himself/herself/themseves’ *en* ‘it’, *ens* ‘us’, *us* ‘you (pl.)’. These variants are also used, together with the nearest etymological forms, in Northern Catalan and Balearic.

eight (i/e/ɛ/ə/a/ɔ/o/u), and their distribution, especially with regard to how they derive from Latin long E and short I (e/ə/ɛ/ɐ) and Latin long O and short U (o/u); the reduction of unstressed vowels to five (i/e/a/o/u), four (i/ə/o/u) or three (i/ə/u or i/a/u); the presence of the phoneme /v/ in the consonantal inventory, or its absence through confluence with /b/; the sibilant inventory, with or without /ʒ/, and with or without other voiced elements (/z/, /dz/ and /dʒ/); the realisation of prepalatal sibilants in initial and postconsonantal position, affricate or fricative ([tʃ]/[ʃ] and [dʒ]/[ʒ]); the realisation of the fricative prepalatal sibilants in final and intervocalic position, with or without a preceding yod ([jʃ]/[ʃ] and [jʒ]/[ʒ]); the evolution of Latin groups -L[j]-, -C'L- and -G'L- either to [ʎ], like L- and -LL-, or to [j]; the treatment of final -r, which can be maintained or lost; and finally the treatment of final consonantal groups of nasal or liquid plus occlusive, which can either be kept or simplified through the occlusive loss.

At the morphosyntactic level, variables include the definite article, coming from ĪPSU or from ĪLLU and, in this latter case, in the classical form *lo* or in the modern form *el*; the personal article, which can either be non-existent, coincident with the definite article or derived from DŌMĪNU; the possessives, in their etymological forms, like *mia*, or in analogical forms, like *meua* and *meva*; the demonstratives, with either a two- or a three-degree system and with either simple forms, like *este*, and *eixe* or reinforced forms, like *aquest*; the clitic pronouns, either in their nearest etymological forms, like *me* < ME, *mos* < *nos* < NOS, or in the more evolved forms, like *em*, *ens*; the inchoative increment from Latin ĒSC, with variations in the tonic vowel (i or ə/ɛ/ɐ/e) and in the maintenance or loss of the velar consonant in the first person present indicative tense ending as well as in the subjunctive forms (-esc- or -isc-/-i/[ʃ]c- versus -eix- or -ix-); the marker of present subjunctive forms, either with the vowels *a*, *e* or *o* (*que jo perda*, *que tu perdes/que tu perdos*) or with the vowel *i* (*que jo perdi*, *que tu perdis*); the first person present indicative tense ending of first conjugation verbs (*cant*, *cante*, *canti*, *cant[o]/cant[u]*) and of other conjugations (*perd/per[k]/perdi/perd[o]/perd[u]*); the evolution of Latin third person verbal forms ending in -AT to [a] (*canta*, *cantava*, *cantaria*) or to [e] (*cant[e]*, *cantav[e]*, *cantari[e]*); and finally the marker of imperfect subjunctive forms, with either a sibilant (*cantés*, *perdés*, *sentís*) or a rhotic (*cantara*, *perdera*, *sentira*).

Despite all these differences, by many standards Catalan is a language with fairly modest dialectal variation. This has made it possible to construct a Standard Catalan that accommodates all of these dialects, with a strongly unified set of orthographic norms (Segarra 1985) which tends to be based on the more conservative dialects – to wit, the distinction between *a/e* and *o/u* in unstressed vowels, the maintenance of *v*, endings with -r, -nt, -lt, etc. – but in which more than one morphosyntactic and lexical

31 In 1989 the journal *COM ensenyar català als adults* published a very useful supplement by Isidor Marí with the assistance of Joan Veny, Josep Moran, Joan M. Romaní and Jordi Bañeres which consisted of several very useful maps detailing the main linguistic variables affecting the dialectal diversity of Catalan as well as other maps showing the distribution of various historical and sociolinguistic data.

variant for the same form have occasionally been granted equal recognition (*GIEC*). Such alternative variants are typically pairs like *meva/meua* ‘my, mine’, *serveix/servix* ‘serves’, *mirall/espill* ‘mirror’, but sporadically more than two are accepted, like *canto/cant/cante/canti* ‘(I) sing’ (⌈14.1 The Language Reform, the Institut d’Estudis Catalans and the Work of Pompeu Fabra). Obviously, the creation of a supradialectal standard was most necessary and beneficial for language use in more formal contexts such as education, the media, public administration, science and literature. Nonetheless, its presence has inevitably had consequences for the original dialects, since any dialectal forms not officially included in the standard (or perceived as not to be included by the population) acquire negative connotations, which hastens their fall from currency into disuse (Sistac 1998; Massanell i Messalles 2012).

5 Bibliography

- Adam Aulinas, Montserrat (2006), *El català septentrional de transició: nova visió des de la morfologia*, Barcelona, Institut d’Estudis Catalans.
- Alarcos Llorach, Emilio (1960), *La constitución del vocalismo catalán*, in: *Studia Philologica. Homenaje a Dámaso Alonso*, vol. 1, Madrid, Gredos, 35–49.
- Alcover, Antoni Maria (1902–1903), *Variedats dialectals que presenta’l català. ¿S’han d’esvaïr, reduïntles a la uniformitat? ¿Hem de centralisar la llengua?*, Bolletí del Diccionari de la Llengua Catalana 1, 149–158.
- Alcover, Antoni Maria (1909), *El català davant els filòlegs estranjers*, Bolletí del Diccionari de la Llengua Catalana 4, 194–304.
- Alcover, Antoni Maria (1918–1919), *Dialectes catalans i llur extensió i característiques principals*, Bolletí del Diccionari de la Llengua Catalana 10, 519–522.
- Alcover, Antoni Maria/Moll, Francesc de Borja (1926), *Introducció*, in: *Diccionari català-valencià-balear*, vol. 1/1, I–LXXI, Palma de Mallorca, Imprenta de Mn. Alcover.
- Alcover, Antoni Maria/Moll, Francesc de Borja (1929–1932), *La flexió verbal en els dialectes catalans*, Anuari de l’Oficina Romànica de Lingüística i Literatura 2, 73–184; 3, 73–168; 4, 9–104; 5, 9–72, <http://alcover.iec.cat/> (last accessed 08.04.2017).
- Alcover, Antoni Maria/Moll, Francesc de Borja (1930), *Diccionari català-valencià-balear*, vol. 1, Palma de Mallorca, Imprenta de Mn. Alcover.
- Alcover, Antoni Maria/Moll, Francesc de Borja (1950–1968), *Diccionari català-valencià-balear*, 10 vol., Palma, Moll, <http://dcvb.iecat.net/> (last accessed 08.04.2017).
- Alsina, Àlex (1988–1989), *Un aspecte de la morfologia històrica catalana: la primera persona singular del present d’indicatiu*, Llengua & Literatura 3, 89–119.
- Badia i Margarit, Antoni Maria (1951), *Gramàtica històrica catalana*, Barcelona, Noguer.
- Badia i Margarit, Antoni Maria (1981), *Gramàtica històrica catalana*, València, Tres i Quatre.
- Badia i Margarit, Antoni Maria (1988), *Sons i fonemes de la llengua catalana*, Barcelona, Publicacions de la Universitat de Barcelona.
- Barnils, Pere (1919), *Dialectes catalans*, Butlletí de Dialectologia Catalana 7, 1–10.
- Baylac-Ferrer, Alà (2009), *Catalunya Nord. Societat i identitat. Reflexions, vivències i panorama català*, Canet de Rosselló, Trabucaire.
- Beltran Calvo, Vicent (2011), *Estudi geolingüístic dels parlars de la Marina Alta. L’empremta mallorquina*, Ondara/Pedreguer, MACMA/IECMA.

- Beltran Calvo, Vicent/Herrero Lloret, Teresa (2011a), *Estudi geolingüístic dels parlars de la Marina Baixa. L'empremta mallorquina*, Ondara/Pedreguer, MACMA/IECMA.
- Beltran Calvo, Vicent/Herrero Lloret, Teresa (2011b), *Atles Lingüístic de la Marina (ALMAR)*, Ondara/Pedreguer, MACMA/IECMA.
- Blasco Ferrer, Eduard (1985), *Les desinències de la primera persona del present d'indicatiu en català i en occità. Estudi diacrònic de morfosintaxi gal-loromànica*, in: *Estudis de llengua i literatura catalanes X, Miscel·lània Badia Margarit*, vol. 2, Barcelona, Publicacions de l'Abadia de Montserrat, 37–87.
- Bover i Font, August (2002), *L'Alguer, confí de confins*, Quaderns d'Italià 7, 111–116.
- Carrera i Sabaté, Josefina (2002), *Escola catalana i variació fonètica. Una evolució del vocalisme àton a Alguairé i a Lleida*, Lleida, Pagès.
- Casanova, Emili (1989), *Gramàtica històrica i història de la llengua. A propòsit de l'evolució de la desinència de 1a persona del present d'indicatiu*, in: *Segon Congrés Internacional de la Llengua Catalana, VIII, Àrea 7. Història de la llengua*, València, Institut de Filologia Valenciana, 343–357.
- Casanova, Emili (1993), *El català dins la Romània: a propòsit del doblat incoatiu esc/isc*, in: *Actas do XIX Congreso Internacional de Lingüística e Filología Románicas*, vol. 5, A Coruña, Fundación Pedro Barrié de la Maza, 293–315.
- Casanova, Emili (2001), *La frontera lingüística castellano-catalana en el País Valenciano*, Revista de Filología Románica 18, 213–260.
- Casanova Solanes, Eugeni (2016), *Els gitanos catalans de França. Llengua, cultura i itineraris de la gran diàspora*, Lleida, Pagès.
- Colomina i Castanyer, Jordi (1985), *L'alacantí. Un estudi sobre la variació lingüística*, Alacant, Institut d'Estudis Juan Gil-Albert.
- Colomina i Castanyer, Jordi (1999), *Dialectologia catalana. Introducció i guia bibliogràfica*, Alacant, Universitat d'Alacant.
- COM *ensenyar català als adults. Suplement 6: Mapes per a l'estudi de la llengua catalana* (1989), Barcelona, Departament de Cultura, Generalitat de Catalunya.
- Conde y Delgado de Molina, Rafael (1994), *Il ripopolamento catalano di Alghero*, in: Antonello Mattone/Piero Sanna (edd.), *Alghero, la Catalogna, il Mediterraneo: storia di una città e di una minoranza catalana in Italia (XIV–XX secolo)*, Sassari, Gallizzi.
- Corbera Pou, Jaume (ed.) (2003), *Arxiu audiovisual dels dialectes catalans de les Illes Balears. Antologia*, Palma, Universitat de les Illes Balears: Càtedra Alcover-Moll-Villangómez.
- Corbera, Jaume/Domènech, Francesc Xavier/Mayoral, Joan (²2000, ¹1997), *L'Alguer: imatge i paraula*, Palma, Universitat de les Illes Balears/Obra Cultural de l'Alguer.
- Coromines, Joan (1954), *El que s'ha de saber de la llengua catalana*, Palma, Moll.
- Coromines, Joan (1971), *Lleures i converses d'un filòleg*, Barcelona, Club Editor.
- Coromines, Joan (1980–2001), *Diccionari etimològic i complementari de la llengua catalana*, 10 vol., Barcelona, Curial/la Caixa.
- Costa, Georges (1977), *Del canvi de “o” tancada tònica en “u” al català del Nord*, Revue de linguistique romane 41, 293–301.
- Escudero, Jean-Paul (1999), *Contribució a l'estudi del canvi de la [ó] tònica tancada en [ú] al Vallespir*, in: *Actes del XXII Col·loqui de la Societat d'Onomàstica (Vilafranca del Penedès, 1996)*, Barcelona, Societat d'Onomàstica, 53–58.
- Escudero, Jean-Paul (2004), *Les gitans catalans et leur langue. Une étude réalisée à Perpignan*, Péronnas, Éditions de la Tour Gile.
- Fabra, Pompeu (1906), *Les “e” toniques du catalan*, Revue Hispanique 15/47–48, 9–23.
- Ferrando Francés, Antoni (1989), *La formació històrica del valencià*, in: Antoni Ferrando (ed.), *Segon Congrés Internacional de la Llengua Catalana, VIII, Àrea 7. Història de la llengua*, València, Institut de Filologia Valenciana, 399–428.

- Fouché, Pierre (1924), *Phonétique historique du roussillonnais*, Toulouse, Privat. Facsimile edition: (1980), Geneva, Slatkine Reprints.
- GIEC = Institut d'Estudis Catalans (2016), *Gramàtica de la llengua catalana*, Barcelona, Institut d'Estudis Catalans.
- Gili i Gaya, Samuel (1931), *Estudi fonètic del parlar de Lleida. Tirada aparte de la Miscelànea Filològica dedicada a D. Antonio Ma. Alcover*, Palma de Mallorca, Círculo de Estudios.
- Gimeno Betí, Lluís (1997), *Atles lingüístic de la diòcesi de Tortosa*, Barcelona, Institut d'Estudis Catalans.
- Gómez Duran, Gemma (2016), *Gramàtica del català rossellonès*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Gran Enciclopèdia Catalana (1973), vol. 4, Barcelona, Enciclopèdia Catalana.
- Griera, Antoni (1931), *Gramàtica històrica del català antic*, Barcelona, Institució Patxot.
- Guinot Rodríguez, Enric (1999), *Els fundadors del Regne de València. Repoblament, antroponímia i llengua en la València medieval*, València, Tres i Quatre.
- Gulsoy, Joseph (1993, ¹1976), *El desenvolupament de les formes de subjuntiu present en català*, in: *Estudis de gramàtica històrica*, València/Barcelona, Institut Universitari de Filologia Valenciana/ Publicacions de l'Abadia de Montserrat, 377–419.
- Gulsoy, Joseph (1993, ¹1987), *La “-o” de la primera persona de l'indicatiu present en català*, in: *Estudis de gramàtica històrica*, València/Barcelona, Institut Universitari de Filologia Valenciana/ Publicacions de l'Abadia de Montserrat, 421–448.
- Hadwiger, Johann (1905), *Sprachgrenzen und Grenzmundarten des Valencianischen*, Zeitschrift für romanische Philologie 29/6, 712–731.
- Hilty, Gerold (1980), *Alguns aspectes de l'evolució històrica de la “e” i la “è” en català*, in: *Estudis de llengua i literatura catalanes oferts a R. Aramon i Serra en el seu setantè aniversari*, vol. 2, Barcelona, Curial, 231–237.
- Kuen, Heinrich (1932, 1934), *El dialecto de Alguer y su posición en la historia de la lengua catalana*, Anuari de l'Oficina Romànica de Lingüística i Literatura 5, 121–177; 7, 41–112.
- Limorti, Ester/Quintana, Artur (2012), *El Carxe: recull de literatura popular valenciana de Múrcia*, Barcelona, Institut d'Estudis Catalans.
- Martines, Josep (2002), *L'aragonès i el lèxic valencià. Una aproximació*, Caplletra 32, 157–201.
- Martines, Josep (2009), *El contacte del català amb la llengua dels aragonesos al segle XIII al País Valencià: influència sobre el lèxic*, Caplletra 46, 61–88.
- Massanell i Messalles, Mar (2012), *“Feva temps que no diva tants verbs!” Manteniment i transformació de paradigmes verbals en el català nord-occidental del tombant de segle*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Massanell i Messalles, Mar (2016), *Sincretisme i canvi morfològic: evolució de la terminació verbal en -AT en català*, Caplletra 61, 165–209.
- Milà i Fontanals, Manuel (1861), *De los trovadores en España. Estudio de lengua y poesía provenzal*, Barcelona, Librería de Joaquín Verdguer.
- Moll, Francesc de Borja (1932), *Estudi fonètic i lexical del dialecte de Ciutadella*, in: *Miscelànea filològica dedicada a don Antonio M. Alcover*, Palma de Mallorca, 397–460. Reproduced in: Randa 8 (1979), 5–48.
- Moll, Francesc de Borja (1980), *El parlar de Mallorca*, Barcelona, Barcino.
- Moll, Francesc de Borja (1991), *Gramàtica històrica catalana*, València, Universitat de València.
- Molla, Guillem (2003), *El català a Andorra: tota una lluita*, *Ianua* 4, 73–90.
- Monclús i Esteban, Joaquim (2014), *La Franja de Ponent: aspectes històrics i jurídics*, Barcelona, Institut d'Estudis Catalans.
- Navarro Gómez, Pere (2017), *La dialectologia catalana a la Universitat Rovira i Virgili. Els microatles lingüístics de les comarques de l'Ebre*, *Estudis Romànics* 39, 361–374.

- Oliver Barreda, Anna (2014), "Jo, parlar, parla normal". *La realització [a] de la primera persona del singular del present d'indicatiu dels verbs de la primera conjugació a quatre localitats de la comarca de l'Alcalatén*, unpublished dissertation, Universitat Autònoma de Barcelona.
- Par, Anfós (1930), *La desinència -o del indicatiu present*, Anuari de l'Oficina Romànica de Lingüística i Literatura 3, 169–176.
- Pérez Saldanya, Manuel (1998), *Del llatí al català. Morfosintaxi verbal històrica*, València, Universitat de València.
- Pons Moll, Clàudia (2007), *La teoria de l'optimitat. Una introducció aplicada al català de les Illes Balears*, Barcelona/Maó, Publicacions de l'Abadia de Montserrat/Institut Menorquí d'Estudis.
- Ponsoda Sanmartín, Joan J. (1996), *El català i l'aragonés en els inicis del Regne de València segons el Llibre de Cort de Justícia de Cocentaina (1269–1295)*, Alcoi, Marfil.
- Rasico, Philip (1982), *Estudis sobre la fonologia del català preliterari*, Barcelona, Curial/Publicacions de l'Abadia de Montserrat.
- Recasens, Daniel (1990–1991), *Tendències fonètiques i classificació dialectal al domini lingüístic català*, Llengua & Literatura 4, 277–310.
- Recasens, Daniel (2017), *Fonètica històrica del català*, Barcelona, Institut d'Estudis Catalans.
- Rigau, Gemma (1992), *Aspects of Catalan Dialectal Syntax. Unaccusative Sentences and Related Structures*, Journal of Hispanic Research 2/2, 315–329.
- Rigau, Gemma (1997), *Locative Sentences and Related Constructions in Catalan: "ésser/haver" Alternation*, in: Amaya Mendikoetxea/Myriam Uribe-Etxebarria (edd.), *Theoretical Issues at the Morphology-Syntax Interface*, Bilbao/Donostia, Euskal Herriko Unibertsitatea/Gipuzkoa Foru Aldundia, 395–421.
- Sardà, Anna/Guiter, Enric (1975), *L'"Atlas Lingüístic de Catalunya" i la fragmentació del català*, Miscellanea Barcinonensia 40, 93–112.
- Scala, Luca (2003), *Català de l'Alguer: criteris de llengua escrita*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Segarra, Mila (1985), *Història de l'ortografia catalana*, Barcelona, Empúries.
- Sistac, Ramon (1998), *El català d'Àneu. Reflexions a l'entorn dels dialectes contemporanis*, Esterri d'Àneu, Consell Cultural de les Valls d'Àneu.
- Sistac, Ramon (2000), *El català nord-occidental i la codificació fabriana*, in: Joan A. Argenter (ed.), *Simposi Pompeu Fabra*, Barcelona, Institut d'Estudis Catalans, 227–234.
- Solà, Joan (1987), *Qüestions controvertides de sintaxi catalana*, Barcelona, Edicions 62.
- Solans, Esperança (1996), *Estudi fonètic sobre la parla de Lleida*, Lleida, Institut d'Estudis Ilerdencs.
- Torres i Torres, Marià (1983), *Aspectes del vocalisme tònic eivissenc*, Eivissa 14, 22–23.
- Vallcorba i Rocosa, Jaume (2011, ¹1987), *Sobre les combinacions binàries de pronoms febles*, in: *Obra gramatical i lingüística completa*, vol. 2, Barcelona, Publicacions de l'Abadia de Montserrat, 273–289.
- Veny, Joan (1980), *Sobre els occitanismes del rossellonès*, in: *Actes del Cinquè Col·loqui Internacional de Llengua i Literatura Catalanes*, Barcelona, Publicacions de l'Abadia de Montserrat, 441–494.
- Veny, Joan (²1984, ¹1978), *Estudis de geolingüística catalana*, Barcelona, Edicions 62.
- Veny, Joan (¹²1998, ¹1978), *Els parlars catalans*, Palma de Mallorca, Moll.
- Veny, Joan (1999), *Aproximació al dialecte eivissenc*, Palma de Mallorca, Moll.
- Veny, Joan (2007–2017), *Petit Atlas Lingüístic del Domini Català*, Barcelona, Institut d'Estudis Catalans, 6 vol. published to date.
- Veny, Joan (2015), *Català occidental / català oriental, encara*, Estudis Romànics 37, 31–65.
DOI: 10.2436/20.2500.01.166.
- Veny, Joan/Massanell i Messalles, Mar (2015), *Dialectologia catalana. Aproximació pràctica als parlars catalans*, Barcelona/Alacant/València, Universitat de Barcelona/Universitat d'Alacant/Universitat de València.

- Veny, Joan/Pons, Lúdia (2001–2018), *Atles Lingüístic del Domini Català*, Barcelona, Institut d'Estudis Catalans, 9 vol.
- Wheeler, Max W. (1984), *La conjugació valenciana: geografia, diacronia i psicologia*, in: *Miscel·lània Sanchis Guarner*, vol. 1, València, Publicacions de la Universitat de València, 409–419.
- Wheeler, Max W. (1995), *La primera persona del present d'indicatiu, pot haver-hi més a dir-ne?*, in: *Estudis de lingüística i filologia oferts a Antoni M. Badia i Margarit*, vol. 2, Barcelona, Publicacions de l'Abadia de Montserrat, 411–425.
- Wheeler, Max W. (2007, ¹1993), *El canvi flexional: els verbs en el català nord-occidental*, in: *Morfologia i fonologia catalana i romànica: estudis diacrònics*, Alacant/Barcelona, Institut Interuniversitari de Filologia Valenciana/Publicacions de l'Abadia de Montserrat. Catalan translation of: *Changing Inflection: Verbs in North West Catalan*, in: David Mackenzie/Ian Michael (edd.), *Hispanic Linguistic Studies in honour of F. W. Hodcroft*, Oxford, Dolphin, 1993, 171–206.

Miquel Àngel Pradilla Cardona

8.2 Social and Functional Variation in Catalan

Abstract: The debate between unity and diversity is alive and well in the Catalan sociolinguistic tradition. This tradition has tackled historical and regional linguistic variation with considerable diligence, but the same cannot be said of social and functional variation. Even so, the work done by sociolinguistics and pragmatics has led to a field of research of great interest.

The aim of this chapter is to characterise the variation within a language community that has internal administrative divisions, interregional communicative dynamics with serious shortcomings and degrees of language minoritization that vary with the region. In the final analysis, the debate on establishing the norm and modelling the standard variety reveals a normative conflict set against a background of doubts about the scope of the language community.

Keywords: sociolinguistics, pragmatics, language variation, standard variety, norm

1 Introduction

As a language enmeshed in a developed cultural and economic environment, Catalan is gradually adapting its linguistic structure to the complexity of the communicative requirements of a language community in the 21st century. The debate between unity and diversity is part of the multifactorial characterisation of a linguistic code that needs to satisfy the requirements of a wide variety of users, who, at the same time, have to be able to cope with a wide variety of communicative situations.

Variation must be studied from two complementary angles: the personal (dialectal) and the functional. In the tradition of research on the Catalan language neither social nor situational variation has been widely touched upon. Sociolinguistics has taken a variety of approaches. On the one hand, Labovian variationism has provided us with an inventory of research that is modest but of great interest; on the other, Anglo-Saxon systemic-functional grammar has spawned a series of sociolinguistic approaches by such authors as Lluís López del Castillo, Isidor Marí, Xavier Lamuela, Lluís Payrató, etc., while sociolinguistics, pragmatics and the ethnography of communication have also focused on the analysis of language use in context.

The sociolinguistic gaze is particularly apt for the study of this duality because of the shortcomings of the specific communicative dynamics. In other words, the Catalan language is generally regarded as a minoritized language because there is a major crisis in language use. We find ourselves, then, in a situation in which it is extremely difficult to establish a language repertoire that is fully accepted by all users of the

<https://doi.org/10.1515/9783110450408-015>

language. What is more, linguistic minoritization manifests itself in different intensities in the different administrative regions that make up the language community. Administrative fragmentation and a history full of disagreements, in turn, encourage the creation of speech communities in which the Catalan language defines its functionality according to its weight with respect to the other codes that make up (an) increasingly complex multilingual linguistic repertoire(s). This segmentation of communicative and referential areas lies at the root of identitarian dynamics that do not favour common norms of usage and shared formal language models. Having said that, nobody should be surprised by the current questioning of the effective existence of the Catalan language community, a concept that goes beyond the strict definition of a territorial unit with a shared historical language. The failure to meet some of the criteria that define community sociolinguistic behaviour requires deep reflection.

In conditions such as these, the precarious nature of the studies available to us must have something to do with the difficulties of characterising certain social and functional varieties, which Catalan language researchers have not focused on as much as historical and geographical aspects. Even when these studies have taken place within the discipline of sociolinguistics, the preference of scholars for sociological analysis has tended to minimise their contribution.

In the current chapter, the sociolinguistics of variation has been given a more central role in the selection of research presented. The study begins by characterising the linguistic variation according to the guidelines of the Anglo-Saxon functionalist doctrine adapted by the Catalan tradition. This is a framework of considerable interest since it can shed a great deal of light on language modelling. We have also felt it necessary to make some reference to the standard variety of language and we have done so by positioning ourselves within the framework of the process of standardisation, a sequence of phases in which the normative code plays a major role. We then close our contribution to the social and functional variation of the Catalan language by sketching out a brief characterisation of the complexity of the current dispute over the monocentric or pluricentric approach to prescription.

2 The characterisation of linguistic variation

Since time immemorial, the study of *interlinguistic variation* – that is to say, the individual or comparative characterisation of different language systems – has attracted the unwavering interest of linguistics. By contrast, linguistics has had real difficulty in accepting *intralinguistic variation* as intrinsic to the very existence of languages. It should be said that while traditional linguistics has dealt with historical and geographical variation, modern linguistics has not felt comfortable in recognising it as a legitimate object of study.

New theoretical models have emerged highlighting the importance of studying intralinguistic variation as part of the general theory of language and they have

spurred researchers on in their efforts to tackle functional complexity. Sociolinguistics and pragmatics, disciplines with considerable overlap, have been making gradual progress in their attempts to define their paradigmatic status.

In terms of the Catalan language, the approach to the compartmentalisation of linguistic variation that has been most widely accepted is the one taken by Halliday/McIntosh/Stevens (1964), which gave rise to systemic-functional grammar and which was revised on various occasions by Halliday (1976) and Gregory/Carroll (1978). We shall refer briefly to this approach below.

Cassany/Marí (1990) suggest two different approaches to linguistic variation in Catalan: associating the features of texts to the origin of speakers or users, and correlating texts with the situations of use in which they take place. The first approach, which deals with *dialects* or *dialectal varieties*, assesses variation according to speakers' geographical origin, social characteristics or time period. The second approach, which deals with *registers* or *functional varieties*, assesses situational varieties; it does not consider users as a group, but rather looks at the diversity of individual uses. The two varieties are two sides of the same coin and are inextricably part of any communicative event. They are, then, two complementary varieties that, when combined, serve to define the variable matrix of all linguistic productions.

Dialectal varieties can be of three different types: *geographic* (traditionally known as dialects), *historical* and *social*. Geographical varieties show how language behaves according to the geolinguistic status of its speakers (for example, Western and Eastern Catalan) (↗8.1 Dialects); historical varieties attribute language to the diachronic stage (preliterary, medieval and contemporary Catalan) (↗10 Early Medieval Catalan; ↗11 The Growth and Expansion of Catalan (1213–1516); ↗12 The Origins of Modern Catalan: Cultural and Linguistic Evolution); and social varieties define linguistic behaviours associated with numerous variables. According to Payrató (1998, 15), these latter varieties are 'cultural varieties, varieties associated with distinctions of social class, ethnic varieties (the gypsy communities in Barcelona and Perpignan), 'abnormal' varieties (speakers afflicted with a language pathology or who speak the language as a second language), gender varieties and varieties belonging to closed and marginal groups (slang) or, at the other end of the scale, more open and inclusive groups (standard varieties).'

The standard variety of the language has had a little trouble finding its position in the above scheme. The fact that it has a clear situational component (it is the variety used in formal communicative situations) and that it aims to neutralise dialectal variation (it is the variety that is least marked by geographical, social and historical features) cannot conceal the fact that it is the dialectal variety of a group of special users, consisting of the totality of the members of the language community.

Below we shall discuss this further but, from the outset, we would like to stress some points that are of particular significance in determining the standard variety of Catalan. The first of these is its internal variability: as well as the variability defined by the situation of use (formal registers), the standard variety also tends to include (albeit

in a somewhat limited fashion) dialectal variation that reflects the identity of the user (regional forms). It is also associated with supradialectal communication so that effective communication can be guaranteed between the various groups of users of the language. Even so, intradialectal communication also tends to have a standard reference, which is more geographically marked. The presence of dialectal forms in the standard variety of the language has a lot to do, on the one hand, with the ease of communication between the various regions of the language community and, on the other, with the identities that make up the linguistic whole. In other words, the consolidation of the standard variety is considerably affected by the limits of the language community. We shall come back to this in the last section of this chapter.

In the introduction to one of the major works in the field, Payrató (1998, 5) defines functional variation in the following way: 'Functional variation – which has also been referred to as *contextual* or *stylistic* – has to do with the differences that are found in a language in the various contexts in which it is used. That is to say, different purposes or functions, interlocutors, channels and topics characteristic of the communicative acts in which human verbal language is used.' It is, then, a modality of situational language associated with particular contexts of use and defined multidimensionally by three factors. Halliday calls these *field*, *mode* and *tenor*, a nomenclature that Catalan sociolinguistics has reformulated with the terms *tema* and *canal* for the first two factors, and the two terms *nivell de formalitat* and *propòsit* for the third.

Field is the object that the interlocutors are referring to, the referential or ideational meaning of the communication. Channel is the medium by which the message is communicated. Tenor refers to the interpersonal meaning of the communicative act and is divided into two complementary aspects: one is the level of formality (personal tenor), which depends on the degree of familiarity or respect between the interlocutors; the other, purpose or intentionality (functional tenor) refers to the effect that is sought with the message.

Hence, registers are the result of a combination of the various factors mentioned. This combination leads to a taxonomy of some considerable complexity. In the determination of the various units of functional analysis, the scale on which the different defining factors are subcategorised plays an important role. This difficulty of establishing discrete units on the functional continuum means that they are particularly complex to study.

The field determines the generic and specific (technical) nature of the communication. It is especially important in defining (thematically unspecified) colloquial and current registers and specialised, scientific or technical registers (technolects). Lexis is the clearest indicator of variations in field and terminology is the central feature of languages of speciality.

Traditionally defined by the written-spoken dichotomy, with the added distinction of prepared and spontaneous uses (planned and unplanned), channel has acquired a new dimension. Namely, a set of text types has emerged that can be placed on the traditional continuum. These texts are half way between oral and written texts,

and have acquired great importance with the advent of the new communication technologies. In the field of the media, the modelling of Catalan mediated colloquial language has been the object of a great many research projects in recent years.

The level of formality of communication is measured on a continuum between the two poles of formal and informal. Categorisation can be extremely complex but, in general terms, three uses are usually defined: informal (colloquial or familiar), neutral (standard) and formal (elevated or solemn). The interlocutors' familiarity with each other and the personal implication of the sender in the communication will have a decisive effect on the choice of the appropriate level of formality.

Finally, purpose or intentionality refers to the functions of language. Jakobson's model ensures that several types are available to us. In our tradition, from the perspective of text and discourse, the proposal made by Adam (1985) (following the work of Werlich 1975) has had considerable impact with its distinction between various text types: narrative (exposition of facts), descriptive (exposition of things), explanatory (exposition of ideas), argumentative (exposition of opinions), injunctive (transmission of orders or instructions), predictive (foretelling of future events), conversational and rhetorical.

Varieties and registers are indissociable realities. The norms governing appropriate choice have become social conventions determined by the sociocultural context. In the case of Catalan, one of the clearest signs of the different degrees of minoritization to which the language is subjected in the places it is spoken is the instability of the *functional repertoire*. In conjunction with *communicative competence*, this concept, which comes from the ethnography of communication (Gumperz/Hymes 1972), is particularly important for characterizing the multiplicity of communicative situations generated by the multilingual environment, territorial atomization and the communicative weakness of the historical language in comparison to Spanish (in the territories belonging to Spain and Andorra), French (in Northern Catalonia and Andorra) and Italian (in the Sardinian city of l'Alguer). At the present time, the functional repertoire of Catalan speakers is multilingual and no consensus has been reached on intralinguistic functional adaptation in the language area as a whole.

In the paragraphs above we have discussed the categorisation of linguistic variation in terms of *varieties* and *registers*. This is the approach that has largely been adopted by the Catalan sociolinguistic tradition, particularly in the debate raised by the definition of the standard, which, as we have seen and should not lose sight of, is a standard variety subject to a wide range of restrictions. Even so, these are not the only ways of categorising variation. As we have seen, *text types* are connected to purpose (personal tenor), one of the four characteristic features of registers and they shift the analysis of communication out of the sociocultural environment and focus on the text from an internal, strictly linguistic, point of view. In addition, if we take a sociocultural approach to communication, we can categorise variation according to *genres*, some of which are associated with daily life (greetings, colloquial conversation) and others with specific activities (courtroom trials, sermons), literary produc-

tion (jokes, riddles), etc. Pragmatics and discourse analysis, two manifestations of the same discipline whose boundaries with (asocial) linguistics are somewhat blurred, and sociolinguistics have dealt with the study of the interaction between language use and communication. Hence, discursive styles, the language of politicians and the media, the construction of identity in discourse, the pragmatic aspects of translation, language models, pragmatististics, phraseology, deixis, modalization, politeness, and markers and connectors have become the foci of researchers working on Catalan. A detailed analysis of the research done can be found in Alturo (2011) and Payrató/Cots (2011).

3 The social and functional view of variation

3.1 The “socialisation” of linguistics

Both traditional and modern linguistics have often looked down on or outright ignored *use* as a subject of study. As a result of this lack of attention, language variability, either for operational methodological reasons or for a firm theoretical standpoint in favour of abstract study, has been regarded as something of a nuisance which is impossible to understand and, in the best-case scenario, only of secondary interest.

The importance given to social context – already quite clear in some approaches in general linguistics – from the 1960s onwards has gradually shaped a research field that shares interests with sociology, anthropology, social psychology, ethnomethodology, pragmatics, discourse analysis, conversation analysis, text linguistics, etc. Thus, despite the opposition of William Labov, the term sociolinguistics has been successfully used to label a field of knowledge that is profoundly interdisciplinary but which is determined to find its paradigmatic autonomy.¹ From this perspective, the understanding of a living language is regarded as being more complete if it reveals not only the structural relations of the system but also how it functions as a social means of communication.

Several scholars have attempted to compartmentalise this wide-ranging, catchall discipline that provides a home to all those studies that, in general terms, focus on social and language-related issues. And despite some discrepancies, researchers seem to be reaching some sort of consensus. Using Joshua Fishman’s (1968, 184) terminological proposal of two main blocks (microsociolinguistics and macrosociolinguistics),

1 Labov’s reluctance to accept the term *sociolinguistics* stems from the fact that the label assumes that it is possible to do linguistics without taking into account the social component. It hardly needs to be said that, despite Labov’s disagreement, this ‘possibility’ has been understood (and continues to be understood) in the sense of ‘highly probable’.

the discipline is largely regarded as having three main areas: a) sociology of language; b) ethnography of communication; and c) variational sociolinguistics.

It should be pointed out that in Catalan circles sociolinguistics was first received with a highly specific purpose in mind. The fact is that Catalan sociolinguistics has been (and still is) essentially a sociology of language. Specifically, what is at stake is the survival of the historical language of the territory. Even so, in recent decades, research has branched out into other areas. Obviously, macrosociolinguistics still plays the leading role, but this approach has now been supplemented with a wide range of variationist and ethnographic research studies. This new focus prompted Boix/Payrató (1996) to remark that the sociolinguistics of conflict in the area of Catalan, had given way to a diversity of sociolinguistics. In the last 20 years, then, ethnography and variationism have made important contributions.

On the other hand, as far as sociolinguistics in the strict sense is concerned (that is to say, the sociolinguistics of variation),² the initial clash of interests with modern linguistics – particularly generative linguistics – has evolved towards a sort of theoretical and methodological understanding. In this regard, we have witnessed the emergence of *corpus linguistics*, which has managed to overcome the introspection and judgements regarding acceptability as a methodological tool, and has focused on the crude reality of the linguistic data (↗9 Language Corpora). Thus, the study of variation has attracted new proponents from beyond the traditions of dialectology and sociolinguistics.

3.2 The sociolinguistics of variation

Labov's sociolinguistic model must be regarded as the most successful product to come from the article by Weinreich/Labov/Herzog (1968). The study in question proved to be a qualitative leap forward in the social consideration of language. When he started out, the American linguist drew on the structural model through the teaching of Uriel Weinreich. However, towards the end of the 1960s he became attracted to generative grammar, which he proposed to supplement with a social focus. From this more realistic perspective, the object of study could be observed in all its diversity.

The main theoretical and methodological proposals put forward by the Labovian model can be summarised as follows:

1. Free linguistic variation does not exist. The heterogeneity of language is not arbitrary and chaotic; it follows an order. Linguistic variation is inherent to the

² With the publication in 1995 of a collection of studies edited by Maria Teresa Turell in the book *La sociolingüística de la variació*, this has become the name by which the discipline is known in the Catalan tradition.

system, of which it is a regular and intrinsic property. There is systematic covariation between a linguistic variable and a sociopragmatic independent variable such as socioeconomic class, gender, age, ethnic group and style.

2. The evolution of linguistic changes can be studied in synchrony, in apparent time, by studying ongoing changes in different age groups in a speech community.
3. From the description and explanation of ongoing changes the existence of linguistic variables correlated with sociopragmatic independent variables can be postulated. Sociolinguistic variables can be: a) *indicators* of social distribution; b) *markers* of stylistic differentiation; and c) *stereotypes*, socially conscious markers. All variants have a social significance.
4. The range of variants cannot be appropriately described in terms of qualitative differences. The description must be quantitative because the variants appear with greater *frequency* in the speech – and sometimes in particular styles – of some social groups than in others. For the quantitative study of speech, a whole methodology has been drawn up around the *sociolinguistic interview*, which includes a typology of contextual styles and stresses the importance of acoustic data, particularly the use of the spectrograph.
5. Grammars have three types of rule: a) invariable or categorical, which have no exceptions and are never violated; b) semi-categorical, with violations that can be interpreted; and c) variable. Variable rules point out that for each speaker of the language the rule will be applied with greater frequency-probability in one context than in another. They account for the fact that linguistic variation is a function of sociopragmatic restrictions that determine the probability with which the rule will be applied, and of the linguistic restrictions or morphonological contexts that encourage or prevent the application of the rule.

In various articles (Pradilla 2002; 2003; 2011), I have attempted to characterise the approaches taken by studies on linguistic variation in the Catalan sociolinguistic tradition, ‘transfers and transactions’ in comparison with the orthodox model. These reviews are based on the multidimensional approach to variation taken by Moreno (1990). According to this author, multidimensionality has four vertices – *geolinguistics*, *historical linguistics*, *pragmatics* and *sociolinguistics* – with the possibility of intermediate positions. We shall now go back to the multidimensional approach and discuss the main studies that can be considered to belong to orthodox Labovian variationism. We shall also discuss pragmatic variationist research.

3.3 Studies on variable phenomena in the Catalan language

3.3.1 The variationist orthodoxy: the sociolinguistic vertex

It is well known that Labovian sociolinguistics has been overwhelmingly used in studies on phonic variables. This is also the case in the Catalan tradition. One of the reasons for this dominance is undoubtedly that it adapts more easily to the premise that variation consists of the alternative use of semantically and pragmatically equivalent forms, something that is difficult to guarantee in lexis and, particularly, syntax. This does not mean, however, that the sociolinguistics of variation in Catalan has not engaged in research on morphosyntactic or lexico-semantic variables. Several studies have been published but with a considerably more limited impact.

Another characteristic that we should not lose sight of is the essentially synchronic approach that Labov himself encouraged. When the Labovian model was adopted by Catalan research, the same bias was reproduced despite the existence of an interesting historicist focus (historical sociolinguistics) (Mas 2003) and the recent emergence of studies on variation in *real time*.

The methodological complexity of the research grouped in the theoretical framework of the sociolinguistics of variation forms the interpretative basis of the modest contribution made by research into the Catalan language. In our opinion, any solid research conducted within this discipline must guarantee that the sample of informants is representative, it must be able to account for the covariation of linguistic variants with groups of linguistic, social and stylistic factors by quantifying results, and, finally, it must be able to provide a qualitative perspective to complement the interpretation of cold statistical results. It hardly needs to be said that the optimal framework for research into these dimensions is a bachelor's thesis, a doctoral thesis or a similarly rigorous study. A wide range of research can be classified as exploratory studies because of the rigorous methods used. In this regard, without entering into detail about the methodologies employed, we feel that particular mention should be made of the contributions made by Gimeno/Montoya (1989), Alturo/Turell (1990), Alturo (1995), Pradilla (1993), Carrera (1999) and Ballart (2004). We shall also briefly mention the variationist research in real time carried out by the project coordinated by Maria Teresa Turell (2004) and the work of Montoya (2000), which focuses on language attrition.

- (1) Prepalatal *apitxament* (not voicing voiced phones) in Petrer (Montoya 1989)³

³ The study, as referenced above, was written by Brauli Montoya and can be consulted in: Gimeno/Montoya (1989, 67–95). Hereafter I refer to Montoya (1989) as an independent item in the bibliography.

This study is part of the author's doctoral thesis *De sociolingüística històrica: canvi lingüístic en curs i desplaçament de llengües a l'extrem meridional de la llengua catalana* (1985). Part of the study describes and explains the linguistic change that consists of not voicing the voiced prepalatal affricate, a phenomenon known as *apitxament prepalatal* (fe[dʒ]e, ro[dʒ]a > fe[tʃ]e, ro[tʃ]a). The speech community studied was from Petrer and the real universe of the sample consisted of Catalan-speaking townspeople (86 informants [1.25 %]).⁴

- (2) The pressure of the standard on prepalatal order in Pont de Suert (Alturo/Turell 1990)

This research project stems from Núria Alturo's bachelor thesis *Canvi sociolingüístic al Pont de Suert* (1987). The variable studied is the voiceless prepalatal affricate, a feature of Ribagorça speech, which is gradually being replaced by the fricative and voiced prepalatal affricate (depending on position) of standard central Eastern Catalan ([tʃ]ent, pa[tʃ]ès, me[tʃ]e > [ʒ]ent, pa[ʒ]ès, me[dʒ]e). The real universe of the sample consists of individuals born in Pont de Suert before 1956 (38 informants over the age of 30 at the time of the interview [7.3 %]). The fact that the sample was adapted to coincide with the stages of social transformation in the town excludes the youngest segment of the population.

- (3) Contemporary evolution of prepalatal deaffrication in Benicarló (Pradilla 1993)⁵

This study is the author's doctoral thesis, which deals with a process of linguistic change which is diachronically referred to as prepalatal deaffrication. The variable is linguistically extremely complex and seven variants have been found through acoustic introspection. These variants can be grouped into two main blocks: fricatives with yod segregation (ro[jʒ]a, me[jʒ]e) and without (ro[ʒ]a, me[ʒ]e), regarded as traditional by geolinguistics, and affricates that are voiced (ro[dʒ]a, me[dʒ]e) and partially unvoiced (ro[dʃ]a, me[dʃ]e) (*rodxa, medxe*), regarded as innovative. The speech community, Benicarló (El Baix Maestrat), is in an area of transition between Western Catalan and Valencian (70 informants [0.72 %]). The considerable presence of the voiced affricate variant /ʒ/ in the older generations in conjunction with the deaffricated variants suggests that the dialectological classification of the linguistic area of deaffrication depends on a surface analysis of the articulations that are present there.

⁴ The real universe of the sample is formed by the individuals of the speech community who are the subject of a sociolinguistic research.

⁵ Some of the publications that have stemmed from this thesis are: Pradilla (1995), (1996a), (1996b) and (1997).

(4) Variation of *haver* and *ser* in Pont de Suert (Alturo 1995)

This research project comes from the same source as that of Alturo/Turell (1990). On this occasion the variable studied is of a syntactic nature: the substitution of the traditional auxiliary verb form *haver* with the verb *ser* where it functions as an auxiliary (*havia parlat/era parlat*).

(5) Alternation of A/E in Segrià (Carrera 1999)⁶

For the first time in the Catalan-speaking area a contrastive analysis is made of how the populations of two different towns – one rural (Alguaire) and the other urban (Lleida) – deal with a particular variable phenomenon. The two speech communities belong to the northwestern geolect (or *lleidatà*). The real universe of the sample consists of Catalan speakers (36 informants from Alguaire [6 %]; 36 informants from Lleida [0.05 %]). An important novelty of the study is that the population studied were children aged between 3 and 5 years old. The study of this age range makes it possible to compare the pronunciations of school and pre-school children. The variable rule aims to describe, explain and predict the presence of the innovative variant [e] and the traditional variant [a] in: the absolute initial pre-tonic vowel written *e* (*encara*); non-absolute initial or medial pre-tonic vowel written *e* (*sencer*); and non-absolute initial or medial pre-tonic vowel written *a* (*calendari*).

(6) The *xava* sociolect in Barcelona (Ballart 2004)⁷

This doctoral thesis studies the *xava* accent, a Barcelona sociolect with phonic features associated with interference from Spanish. The most representative feature is the opening of the neutral vowel (*par[ə]/par[a]*), but also important are the voiceless alveolar and prepalatal fricatives (*ca[z]a/ca[s]a* and *bo[ç]a/bo[ʃ]a*). The study focuses on the communicative behaviour of adolescents, both boys and girls, from Barcelona. Methodologically speaking this research is noteworthy in the Catalan tradition because it asked the informants to declare their L1 (Catalan or Spanish).

(7) Variationist studies in real time in Catalan (Turell, principal investigator 2004)⁸

⁶ Some of the publications that have been derived from this thesis are: Carrera (2001a), (2001b), (2002a), (2002b) and (2002c).

⁷ A related publication can be found in Ballart (2013).

⁸ *El efecto de los factores internos y externos en el contacto lingüístico y en los procesos de cambio lingüístico en 4 comunidades catalanohablantes: La Canonja, El Pont de Suert, Benicarló y Petrer* (Projecte HUM2004-0504-C02-01/2005-2009).

This project has engaged in new research into speech communities that were studied in the 1990s and has incorporated real-time studies into Catalan variationism by replicating previous research several years afterwards. Among other research, new data has been provided on the variables studied in Pont de Suert (Alturo/Turell 1990 and Alturo 1995) and Petrer (Montoya 1989). The results on the situation of the ongoing changes in both speech communities were different; whereas the change in Petrer has continued (Montoya 2006; Verdú 2007a and 2007b), in Pont de Suert it could be said that it has reversed since the younger generations have largely opted to maintain the local varieties, particularly those of a phonetic nature (Suïls et al., 2010).

(8) Language attrition and shrinkage in Alacant (Montoya 2000)

In this study the author initiates a line of work on language attrition and shrinkage. This focus is of particular interest in a context such as the Catalan one in which there is language contact. The work successfully complements the author's research on the interruption of intergenerational language transmission in Alacant (Montoya, 1996) and deals with the erosion or structural disintegration of the recessive or endangered language (Catalan) in the context of convergence with the expanding language (Spanish).

Having made a brief summary of the research as a whole, we shall now move on to discuss some methodological concerns. In terms of the phonic variables studied, particular interest has been shown in palatal consonants, an articulatory series that varies greatly in the Catalan-speaking area. In the speech communities investigated, the Labovian model has been largely deurbanised. With the exception of Barcelona and Lleida, all the other places studied are rural villages (El Pont de Suert, Alguaire) or medium-sized towns (Petrer and Benicarló). In geolinguistic terms, the studies have focused above all on speech communities located in areas of language transition or at some distance from cities that have adopted the standard model. Finally, we should point out that apart from Ballart (2004), all the other studies have selected Catalan speakers as the real universe of the sample even though the communities are multi-lingual.

As far as the starting hypotheses of the research discussed are concerned, it should be noted that linguistic factors are largely regarded as important influences on the changes studied. Therefore, the linguistic characterisation of the variables has attracted a great deal of interest. The other interpretative vertices have been language interference and the Catalan formal model.

In terms of social factors, age is undoubtedly the one that has been most widely studied. Changes in apparent time (generational distribution of variants) have been studied in most of the research mentioned. Except for Ballart (2004), who focuses on adolescents from Barcelona, the variable age has been of particular importance. In the studies by Montoya (1989), Pradilla (1993) and Carrera (1999), the younger age groups have been at the forefront of the ongoing changes. The work by Alturo/Turell (1990)

and Alturo (1995) has shown that, even though the age groups considered (informants older than 30) do not allow such a complete analysis, autochthonous variants, particularly syntactic ones, are more common in the elderly groups. Meanwhile, in terms of the phonic variable, the youngest age group tends to behave in the same way as the oldest.

The factor of sex (gender) has produced a variety of results in the various studies but it is also of considerable importance. Women were observed to be leading the changes in the studies by Montoya (1989), Alturo/Turell (1990), Pradilla (1993), Carrera (1999) and Ballart (2004). On the other hand, Alturo (1995) shows that it is men who were promoting the standard variety. In this respect, it should be pointed out that a great deal of research has confirmed the major role that women play in promoting prestigious varieties of language. In the studies mentioned above the correlation between the gender variable and the other social variables considered reveals even more precise results. In all the studies, change is driven by young women, and in Pradilla (1993) and Carrera (1999) exposure to education is also fundamental. Hence, the adaptation of pronunciation to the graphic representation of the variants analysed (no yod segregation when informants from Benicarló read texts and the phonetic spelling – the variant [e] – of the informants from Alguaire/Lleida) has a lot to do with how well the members of the communities in question understand normative Catalan.

The factor of *social class* deserves special mention, having been studied by Montoya (1989) and Carrera (1999) with negative and positive results, respectively. Alturo/Turell (1990) and Alturo (1995) presented the factor *socioeconomic sector* as an alternative. Pradilla, on the other hand, conducted a separate study of *employment level, relational area* (for retired people, housewives or students) and *level of education*. Meanwhile, Ballart (2004), even though he discussed social class, also took into account the profession of the parents of the Barcelona adolescents interviewed.

As far as these socioeconomic and/or sociocultural factors are concerned, it could be concluded that language changes have been initiated *from the top down* and *from the bottom up*. The former are conscious changes and link the use of prestigious varieties to the higher social classes. This is what happens when pronunciation is adapted to the grapheme *e* in pre-tonic position in Ballart (2004). In Alturo/Turell (1990) and Alturo (1995), the change in the standard fricative articulation and the substitution of *ser* for *haver* in the past perfect is driven by intermediate groups who have a greater degree of social interaction with members of other communities. Pradilla (1993) found that students, technical professionals and the service sector were spearheading the spread of semi-unvoiced affrication. This latter case is an example of bottom-up, unconscious change caused by the widespread use of Spanish in these groups, which leads to a variant that is halfway towards pre-palatal *apitxament*.

The *sociolinguistic interview* is a methodological device used to segment the stylistic continuum according to the extent which speakers are aware of their own

discourse. The only conclusion that can be drawn about the interview is that authors have developed their own personal versions. The study that is most closely adapted to Labov's contexts is Montoya's (1989) and the results did not meet the initial expectations. Pradilla (1993) attempted to create a formality continuum by eliciting narratives, an introductory conversation, a selection of words to translate and the reading of a text. Although there was no formally established model in the Valencian Country at the beginning of the 1990s, the semi-vowel of the traditional variant of yod segregation was not reflected in spelling and was clearly in a minority. Ballart (2004) also noted a greater presence of the standard variety in communicative formality (individual conversation). The other studies mentioned have not assessed this factor. This is the most obvious shortcoming of the studies by Alturo/Turell (1990) and Carrera (1999).

3.3.2 The pragmatic vertex

As we have just seen, the variables studied by the sociolinguistics of variation are not discrete; they are defined in terms of statistical correlation. Even so, variation can also be discrete: the most oft-cited example is the choice of language, though the various phenomena involved in interference, codeswitching or the emergence of interlanguages (pidginisation and creolisation) also need to be borne in mind. Pragmatics, understood to be the adaptation of language use to the sociocultural context, deals with these issues. Likewise, the ethnography of communication has emerged as an alternative to variationism and has a clear connection to pragmatics. With certain anthropological concerns, the ethnography of communication conceives the study of linguistic variation in a human community 'as the analysis of how its members manage the phenomenon of heteroglossia in their daily lives from the basis of an agentive and dialogic conception of human communication' (Argenter 1998, 16). Using a system of units of analysis that go into ever greater detail – communicative act, speech event and speech act – ethnographers aim to correlate language use with the sociocultural context of a speech community. Their task, then, requires speech acts, the associated ways of speaking and the surrounding socio-communicative contexts to be identified.

Now that this initial characterisation has been made, we shall move on to focus on the points of conflict between sociolinguistics and variationism: a) whereas the latter is primarily interested in spontaneous speech ("vernacular"), the former focuses on the formal extreme of the stylistic continuum (that is to say, speech that is often ritualised, aphoristic and marked); b) whereas the Labovian model of dividing up the stylistic continuum is based on the extent to which speakers are aware of their discourse, ethnographers seek natural varieties of language in particular contexts; and c) whereas the object of study in sociolinguistics is approached largely through the preferential use of qualitative techniques such as participant observation, varia-

tionism uses statistical approaches that are not so natural (for example, the sociolinguistic interview).

Various aspects of the ethnography of speech can make interesting improvements to variationist methodology. One of the most fundamental is the achievement of greater equilibrium between qualitative and quantitative analysis. The other main issue on which ethnography has a lot to say is the subcategorisation of functional variation. Since stylistic variation occupies a privileged position in pragmatics, variationism should be able to take advantage of the accumulated experience of numerous research projects. Although the segmentation of the stylistic continuum is still generating considerable debate, pragmatics is the discipline that has best got to grips with it. In the field of Catalan linguistics, Payrató (1998), must be regarded as the most complete reference work available.

It is probably no coincidence, then, that one of the people who made a contribution to this book, Sílvia Romero, focused his doctoral thesis (Romero 2001) along similar lines. The pragmatically based study explains how the use of quantitative methods of sociolinguistic methodology can be successfully used to analyse a communicative situation as fixed as council plenary sessions. The linguistic expression of the local administration has been characterised using Hymes' list of components of speech events (1968, 110–124; 717 *Languages in Contact: A Sociocultural Approach*) and Biber's situational parameters of variation (1994, 40–41).

4 The standardisation process of the Catalan language

4.1 The standardisation of subordinate languages

Lamuela (1994, 137) lists the following features of the structure of subordinate languages: a) excessive non-functional variability, which adds little in terms of language efficiency; b) fragmentation into dialects that to some extent or another are mutually unintelligible; c) no, or limited, means of expression in some areas; d) a reduced or fragmented organisation of registers; e) acceptance of massive interference and the evolution of the dominant language.

Hence, the norm of subordinate languages is defined more imprecisely than that of established languages and it is subject to laxer social control because it is not evident in all uses of formal communication. The restrictions on its use – which are social and geographical as well as functional – are subject to a fragmentary working order that prevents varieties from finding a balance and reduces the systems of connotation. Language interference is considerable and it destabilises the internal norm, while the fragmentation of the system of connotation and stylistic interference contribute to the instability of stylistic norms. This superposition of the dominant

language will end up making its standard forms the reference and directing the processes of language levelling and innovation.

Lamuela (1994, 104–105) defines five fundamental processes in the characterisation of languages in regression: a) decreasing use of the language; b) loss of areas of definition; c) contraction and fragmentation of the stylistic system; d) increasing restrictions on norms of use; e) reduction in the connotations that are features of strongly characterised areas and devaluation of the language.

Having said this, it should be pointed out that the two sequences – structural features and the processes of regression – are closely linked. In fact, the physiognomy of the linguistic structure largely depends on the social processes listed. Therefore, we only have a chance of successfully standardising a subordinate language if the process of standardisation takes place within a determined process of normalisation.

It should be made clear from the outset that the Catalan-speaking diasystem as a whole perfectly matches the list of characteristic features of contexts of linguistic minoritization just mentioned. Even so, it is only fair to point out that the regional differences require individual assessments. From this sectorial viewpoint, it can be seen that there is a scale of linguistic (ab)normality which should not be ignored.

Indeed, according to the *Informe sobre l'estat de la llengua 2015 (Report on the state of the language 2015, Pradilla/Sorolla 2016)* the sociolinguistic situation of the Catalan language shows a wide range of quite different regional dynamics. In this regard, some regions can reasonably expect to achieve a sufficient status of vitality while others reveal ongoing processes of fairly intense language shift. Andorra and Catalonia are the regions where the sociolinguistic indicators are most favourable, meaning that the Catalan language has a certain capacity of attraction. We are referring, in particular, to the favourable difference between an individual's first language and the language of identification, and the vitality of intergenerational language transmission. On the other hand, in Northern Catalonia and l'Alguer, the opposite is true. In between these two groups of regions there are also two other regions in intermediate situations: the Balearic Islands tend to share more features with the first group whereas the Valencian Country tends to resemble the second. In the Aragonese Strip (la Franja in the Autonomous Community of Aragon), where Catalan had established a majority presence in local colloquial communication in a diglossic situation of functional distribution, the latest sociolinguistic indicators have begun to reveal signs for concern.

The overall characterisation of the language community will largely depend on the sociolinguistic evolution of the Balearic Islands and the Valencian Country because these two regions, in conjunction with Catalonia, make up 95 % of the language population. For our present purpose, the assessment of language competence and prevailing usage in these three regions will be fundamental. As far as communicative abilities are concerned, Catalonia and the Balearic Islands go hand in hand. In both cases almost everybody understands the language, more than 80 per cent can speak it and more people can read it than write it. Of the four skills, writing

represents the lowest percentages. The most negative indicators can be found in the Valencian-speaking part of the Valencian Country; one out of every four people claims not to understand the language, just over half claim they can speak and read it, and slightly more than a third claim they can write it.

In terms of usage, the indicators of habitual language (main language of use) in both Catalonia and the Balearic Islands are clearly unfavourable to Catalan in its struggle with Spanish. The difference is 14 in Catalonia and 13 points in the Balearic Islands. In the Valencian Country the difference is even greater (between 25 and 51.5 points depending on the contexts analysed).⁹ Likewise, it should also be pointed out that when people leave the comfort zone of contact with their ethnolinguistic group, that is, when they are unaware of the linguistic identity of their interlocutor, Catalan loses ground in the Balearic Islands and, above all, in the Valencian Country. On the other hand, in Catalonia, as a result of more targeted public policies, Catalan is gaining ground outside the home.

In this respect, we should not lose sight of the fact that Spanish is omnipresent throughout all these regions. It plays an important role in the linguistic repertoire of those people whose first language is Catalan and it is the *lingua franca* of most of the international migrants who arrived around the turn of the millennium. Programmes encouraging migrants to embrace Catalan have only had a significant following in Catalonia as part of a language policy that is committed to promoting the historical and territorial language as the code for inter-group relations. The results, however, have left a great deal to be desired; Catalan has always been learned alongside Spanish but a command of Spanish often means that there is no real need to have a command of Catalan.

Territorial shrinkage, the curtailment of uses, a reduction in the stylistic system, and structural and connotational interference from the dominant language are all factors that place Catalan in a situation of linguistic subordination. An awareness of the fact that the Catalan language is immersed in negative communicative dynamics, which vary in intensity depending on the region, should guide the decisions that are made in the field of corpus planning.

4.2 Establishing the language norm

Ever since it was founded in 1907 the Institut d'Estudis Catalans (IEC) has been responsible for establishing the norm and guiding the process of standardisation throughout the linguistic area. Even so, at present, with the advent of new regional codifying centres, the standardisation of the language is in a state of considerable disarray. It could be said that the setting up of the Acadèmia Valenciana de la Llengua

⁹ The differences are for the contexts of *at home* and *at the megastore*, respectively.

(AVL) (Law 7/1998, of 16 September, on the creation of the Valencian Academy of the Language) initiated new normative-standardisation conditions, which could be called *pluricentric*. These conditions were reinforced with Law 10/2009, of 22 December, on the use, protection and promotion of the languages of Aragon, which envisaged the creation of the Aragonese Academy of Catalan.

The passage of time has altered the trend somewhat. In the case of Aragonese, the successive changes of government (and linguistic and ideological standpoints), first in the autonomous community elections in 2011 and then four years later in the elections of 2015, have meant that none of the announced proposals have been implemented. Meanwhile, as far as Valencia is concerned, the new Academy has become just one more example of the particularism that has been promoted by broad sociopolitical sectors of the Valencian Country. This particularism has undergone a substantial transformation that has prompted the abandonment of traditional secessionism and the encouragement of referential isolationism.

Whatever the case may be, it is clear that the traditional model of monocentric management of the norms governing the Catalan language is now being challenged. The policies and ideologies that prevailed at least up to 2015 prevented consensus on various points and, thus, could not guarantee that the code would function as a single unit.¹⁰ Underlying everything is the debate on the scope of the sociolinguistic community, a debate that has much in common with the aspirations of functionality that all political leaders want for the Catalan language. In this regard, we should recall once again that a reliable process of corpus planning ought to take place in a general framework of language planning and must be in tune with any interventions on language status.

If we look back, around the year 1932 various events took place that helped to establish the normative Catalan grammar sponsored by the IEC, a prescriptive corpus that was based on three works: a) the IEC's *Normes ortogràfiques* (*Orthographic Norms*, 1913), slightly modified in the *Diccionari ortogràfic* (*Orthographic Dictionary*, 1917); b) Pompeu Fabra's *Gramàtica catalana* (1918); and c) the *Diccionari General de la Llengua Catalana*, also by Pompeu Fabra (1932). These events were, on the one hand, the decision by Barcelona's Acadèmia de Bones Lletres to agree to use the new standards, thus reducing all opposition to a small minority; and, on the other, the death of Antoni M. Alcover, which sadly but effectively put an end to anti-Fabrism in the Balearic Islands. The consensus that led to the signing of the Norms of Castelló in 1932 also played an important role. These norms meant that Fabra's doctrine had been accepted in València, and that an important territory had joined the common cultural

10 The autonomous community elections in May 2015 led to changes in the government in Aragon, the Valencian Country and the Balearic Islands. The People's Party, whose language attitudes did little to encourage the restoration of the Catalan language or to address the language community as a whole, gave way to progressive coalitions whose language policies, at least on paper, were more favourable to both objectives and envisaged fluid relations between its component parts.

project with a shared language as its central theme (↗14.1 The Language Reform, the Institut d'Estudis Catalans and the Work of Pompeu Fabra).

The year 1932, then, was fundamental in establishing and spreading the norms that were to govern the whole language community. In short, it was the year that the unity of the language ceased to be merely a point on an electoral programme. The implementation of the language standard throughout the vast territorial domain made the drive for unity much more effective, despite occasional attempts at opposition. The first third of the 20th century, which was concluded by the events of 1932, then, was a period that defined the central themes that would form a part of the process of standardisation.

The normative grammar of the Catalan language was drawn up on the basis of the criteria of diasystematicity and historicity (↗3 Spelling). Thus, the codification model proposed could be described as *monocentric*, essentially revolving around one codifying centre (Institut d'Estudis Catalans), but also compositional and polymorphic (that is to say, with contributions from the various speech varieties and, at the same time, accepting of alternative forms as equivalent solutions). It is the establishment of the orthographic norms – of foremost symbolic importance – that best demonstrates the compositional nature of the process. However, we would like to draw attention to the ideological background that complaints of a prevalence of morphological and lexical options from central eastern Catalan tend to conceal because these complaints ignore that the compositionality clearly favours western varieties (particularly Valencian) in the writing system. Even though the linguistic complaint is well founded in part, the users' identification with or rejection of the formal model is reminiscent of extra-linguistic considerations that have a lot to do with the centrifugal dynamics that have become chronic in the Catalan language community. And as unequivocal proof of this referential fragmentation, we point to the current shift away from the original monocentrism in favour of a pluricentrism adapted to the organisation of the Catalan-speaking territories as autonomous communities within the Spanish state.

Briefly, it could be said that the new Valencian proposal, as part of the arduous and permanent negotiation between the different linguistic and ideological sensibilities that coexist within the AVL, has become a synthesis of the fundamental principles of secessionism (individuation, colloquialisation of the norm and the legitimatisation of linguistic interference) and projects them on the Fabrian codification. It should also be noted that this operation has also given rise to a conflict with the prescriptive corpus of the IEC. The dissonances have become particularly clear in lexical prescription, first with the *Diccionari ortogràfic i de pronunciació del valencià* (*Orthographical and Pronunciation Dictionary of Valencian*, 2006) and more recently with the *Diccionari normatiu valencià* (*Normative Dictionary of Valencian*, 2014).

One of the most worrying questions raised by this model is the confusion generated by the stylistic readjustment it promotes. The publication of the *Gramàtica normativa valenciana* (*Valencian Grammatical Norms*, 2006), which raised forms that had previously been classified as colloquial to the highest levels of formality (unrein-

forced demonstratives – *este, esta, estos, estes* –, an increase in the use of inchoatives – *ix-constituïx, deduïx* –, etc.), the cultural agents who have led the struggle to dignify the language (academic, the education sector and the publishing sector) have been radically challenged. Consensus will not be easy although, with time, institutional pressure and all the mechanisms it has at its disposal may tip the balance.

Meanwhile, the IEC has been redrafting its compositional and polymorphic codifying proposal. In the field of lexis, the publication in 1995 of the *Diccionari de la llengua catalana* (*Dictionary of the Catalan Language*, 1995) substituted Pompeu Fabra's work (the second and latest edition, in 2007, made substantial changes). Very recently, the new *Gramàtica de la llengua catalana* (*Grammar of the Catalan Language*, 2016) added weight to the diasystematic perspective. The new guidelines for the written language reflect the fundamental principles that the IEC had defined for the oral language in its proposals for the phonetic and morphological oral standard at the beginning of the 1990s (IEC 1990; 1992). With a view to opening up the normative grammar, the academy asks language users to agree on a model based on communicative appropriateness. Geographic and functional variation take on a new dimension that should more comfortably align with the normative and standardising discrepancies mentioned above.

4.3 In search of a communicative community

Establishing a standard variety assumes the awareness of the existence of a language community. In the case of Catalan, the long list of territorial disagreements throughout history has led to the perception of individual belonging and an extreme lack of communicative behaviours between regions.

The concepts of language community and speech community may shed some light on our debate (Mas 2008, 95–144). They can be used to describe different spaces of functional and symbolic relations. The largest of these, and also the weakest in terms of the communication among its members, is the language community; and the smaller but communicatively more compact units are the speech communities. The sum of the speech communities constitutes the language community; and the smaller units can be defined using a variety of criteria. Thus, the language community may be the result of the additive segmentation of a wide range of communicative spaces such as town, administrative region, district, etc. These different levels of communicative interaction are associated with different levels of awareness and symbolisation.

The debate about the factors that supposedly guide the establishment of a language community takes on special meaning in the case of Catalan. For reasons of space, we cannot provide here the whole list of characterisations provided by the sociolinguists who have focused on the theoretical aspects of this issue. We shall merely point out that there are two blocks of criteria: one of these blocks deals with interaction among speakers and the other places the emphasis on factors of a sym-

bolic or an affective nature. Whereas the former stresses effective communication among the members of the community, the latter focuses on agreement with the norms and attitudes to the language of the members of the community. Since in situations of plurilingualism, ideological plurality guarantees discrepancies in terms of norms and attitudes to the languages in the linguistic repertoire, we believe that a language community should be defined on the basis of the criterion of communicative interaction, of the effective inter-relation of its members. Thus smaller communication spaces can be constructed that are fully interconnected among themselves.

At present, territorial discrepancies in the standard model of language are a clear indicator for the persistence of identitarian dynamics in the definition of the groups of reference. These dynamics are accentuated by a lack of interterritorial communication. Likewise, the alternative model proposed by the AVL goes further than the IEC's diasystematic proposal. It is an attempt to define those features that set it apart, to highlight the symbolic values of the Valencian variety (identification of the Valencian user with the model) rather than the instrumental (communicative) values.

It is clear that at the present time the normative management of Catalan is closer to a model with a variety of codifying centres and standard varieties than a unitary model that administers diversity by following criteria of supraterritorial communication. The new standardising situation is a further indicator of the failure of symbolic and attitudinal factors and the extraordinary weakness of the interactional factors that define the existence of the language community. The viability of these communicational factors must be the main focus of any process of linguistic planning committed to Catalan language.

5 Bibliography

- Adam, Jean-Michel (1985), *Quels types de textes?*, *Le français dans le monde* 192, 39–43.
- Alturo, Núria. (1995), *La variació d'haver auxiliar al català nord-occidental*, in: Maria Teresa Turell (ed.), *La sociolingüística de la variació*, Barcelona, PPU, 221–255.
- Alturo, Núria (2011), *Pragmàtica i anàlisi del discurs*, *Treballs de Sociolingüística Catalana* 21, 29–41.
- Alturo, Núria/Turell, Maria Teresa (1990), *Linguistic Change in El Pont de Suert. The Study of Variation of /3/*, *Language Variation and Change* 2, 19–30.
- Argenter, Joan A. (1997), *Els fenòmens de la diversitat i la variació lingüístiques, i llur pertinència en l'estudi del llenguatge*, in: Maria Rosa Lloret et al. (edd.), *Anàlisi de la variació lingüística*, Barcelona, PPU, 15–44.
- Argenter, Joan A. (1998), *La variació lingüística: el fet i la seva significació*, *Caplletra* 25, 11–19.
- Ballart, Jordi (2004), *Phonological variation in Barcelona Catalan: the Xava accent*, Ph. D. thesis, Queen Mary College, University of London.
- Ballart, Jordi (2013), *Variació fònica al català de Barcelona*, *Treballs de Sociolingüística Catalana* 23, 133–151.
- Biber, Douglas (1994), *An Analytical Framework for Register Studies*, in: Douglas Biber/Edward Finegan (edd.), *Sociolinguistic Perspectives on Register*, New York/Oxford, Oxford University Press, 31–56.

- Biber, Douglas (1995), *Dimensions of Register Variation. A Cross-Linguistic Comparison*, Cambridge, Cambridge University Press.
- Biber, Douglas/Finegan, Edward (edd.) (1994), *Sociolinguistic Perspectives on Register*, New York/Oxford, Oxford University Press.
- Boix, Emili/Payrató, Lluís (1996), *An Overview of Catalan Sociolinguistics and Pragmatics (1989–1996)*, *Catalan Review* 9, 317–403.
- Carrera, Josefina (1999), *L'alternança A/E al Segrià*, Ph. D. thesis, Universitat de Barcelona.
- Carrera, Josefina (2001a), *Algunes consideracions generals sobre l'anàlisi acústica de [e] i de [a] àtones*, *Estudios de Fonética Experimental* 11, 67–87.
- Carrera, Josefina (2001b), *La normativització del català modifica els hàbits fonètics dels parlants?*, *Llengua & Literatura* 12, 175–199.
- Carrera, Josefina (2002a), *L'estadística en l'anàlisi de la variació fonètica: una aplicació del programa Goldvarb*, *Noves SL*, Primavera 2002, 1–5. <http://cultura.gencat.es/llengcat/noves/hemeroteca/metodologia/carrera> (last accessed 10.04.2018).
- Carrera, Josefina (2002b), *La identitat dels processos de canvi lingüístic. Un cas del lleidatà*, *Estudis del Valencià d'ara*, Actes del IV Simposi de Filologia Valenciana, València, Denes, 33–43.
- Carrera, Josefina (2002c), *Escola catalana i variació fonètica: Una evolució del vocalisme àton a Alguair i a Lleida*, Lleida, Pagès.
- Cassany, Daniel/Marí, Isidor (1990), *Els perfils del català. Varietats i registres de la llengua catalana*, COM ensenyar català als adults 4.
- Fishman, Joshua A. (ed.) (1968), *Readings in the Sociology of Language*, The Hague, Mouton.
- Gimeno, Francisco/Montoya, Brauli (1989), *Sociolingüística*, València, Universitat de València.
- Gregory, Michael/Carroll, Susanne (1978), *Language and Situation: Language varieties and their social contexts*, London/Boston, Routledge & Kegan Paul.
- Gumperz, John J./Hymes, Dell H. (edd.) (1972), *Directions in sociolinguistics. The ethnography of communication*, New York, Holt, Rhinehart & Winston.
- Halliday, M. A.K. (1976), *Language as Social Semiotic. The Social Interpretation of Language and Meaning*, London, Arnold.
- Halliday, M. A.K./McIntosh, August/Stevens, Peter (1964), *The Linguistic Sciences and Language Teaching*, London, Longman.
- Hymes, Dell H. (1968), *The Ethnography of Speaking*, in: Joshua A. Fishman (ed.), *Readings in the Sociology of Language*, The Hague, Mouton, 99–137.
- Hymes, Dell H. (1971), *On Communicative Competence*, Philadelphia, Pennsylvania Press.
- IEC (1990), *Proposta per a un estàndard oral de la llengua catalana. I Fonètica*, Barcelona, Institut d'Estudis Catalans.
- IEC (1992), *Proposta per a un estàndard oral de la llengua catalana. II Morfologia*, Barcelona, Institut d'Estudis Catalans.
- Labov, William (1966), *The social stratification of English in New York City*, Washington D. C., Center for Applied Linguistics.
- Labov, William (1972), *Sociolinguistic patterns*, Philadelphia, University of Pennsylvania Press.
- Lamuela, Xavier (1994), *Estandardització i establiment de llengües*, Barcelona, Edicions 62.
- Mas, Antoni (2003), *La sociolingüística històrica: una alternativa a l'anàlisi del canvi lingüístic*, *Noves SL* (Autumn). http://www.gencat.cat/llengua/noves/noves/hm03tardor/mas1_2.htm (last accessed 10.04.2018).
- Mas, Antoni/Montoya, Brauli (2004), *La sociolingüística de la variació als Països Catalans: estat de la qüestió*, *Caplletra* 37, 243–266.
- Mas, Josep Àngel (2008), *El morfema ideològic. Una anàlisi crítica dels models de llengua valencians*, Benicarló, Onada Edicions.

- Montoya, Brauli (1989), *Estratificació de la variació lingüística a Petrer*, in: Francisco Gimeno/Brauli Montoya, *Sociolingüística*, València, Universitat de València, 67-95.
- Montoya, Brauli (1996), *Alacant: La llengua interrompuda*, València, Denes.
- Montoya, Brauli (2000), *Els alacantins catalanoparlants: Una generació interrompuda*, Barcelona, Institut d'Estudis Catalans.
- Montoya, Brauli (2006), *Enquesta en temps real a Petrer*, presentation at the XIV Col·loqui de l'Associació Internacional de Llengua i Literatura Catalanes, Budapest, Hongria, 4 al 9 de setembre de 2006.
- Moreno, Francisco (1990), *Metodología sociolingüística*, Madrid, Gredos.
- Payrató, Lluís (ed.) (1998), *Oralment: Estudis de variació funcional*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Payrató, Lluís/Cots, Josep (edd.) (2011), *The Pragmatics of Catalan*, Berlin/Boston, De Gruyter Mouton.
- Pradilla, Miquel Àngel (1993), *Variació i canvi lingüístic en curs al català de transició nord-occidental/valencià*, Ph. D. thesis, Tarragona, Universitat Rovira i Virgili.
- Pradilla, Miquel Àngel (1995), *El desafricament prepalatal en el català de transició nord-occidental/valencià*, in: Maria Teresa Turell (ed.), *La sociolingüística de la variació*, Barcelona, PPU, 53-116.
- Pradilla, Miquel Àngel (1996), *La sociolingüística quantitativa i la fonètica experimental: per una complementarietat metodològica*, in: *Actes del I Congrés de Lingüística General. Panorama de la investigació lingüística a l'Estat Espanyol*, València, Universitat de València, 142-150.
- Pradilla, Miquel Àngel (1997a), *Sociolingüística de la variació i actituds interdialectals a l'extrem nord del País Valencià*, in: Mercè Pujol Berché/Fermín Sierra Martínez (edd.), *Las lenguas en la Europa comunitaria II: Las lenguas de minorías* (= Diálogos Hispánicos 19), Leiden, Brill, 97-124.
- Pradilla, Miquel Àngel (1997b), *Sociolingüística quantitativa i anàlisi qualitativa de variables fonetico-fonològiques: a propòsit del desafricament prepalatal*, *Estudios de Fonética Experimental* 8, 207-251.
- Pradilla, Miquel Àngel (1998), *La sociolingüística de la variació*, in: Miquel Àngel Pradilla (ed.), *Ecosistema comunicatiu: Llengua i variació*, Benicarló, Alambor, 15-44.
- Pradilla, Miquel Àngel (2001a), *La sociolingüística de la variació: aproximació metodològica (I)*, *Noves SL* (Winter-Spring). <http://cultura.gencat.es/llengcat/noves/hemeroteca/metodologia.htm> (last accessed: 10.04.2018).
- Pradilla, Miquel Àngel (2001b), *La sociolingüística de la variació. Aproximació metodològica (i II)*, *Noves SL* (Autumn). <http://cultura.gencat.es/llengcat/noves/hemeroteca/metodologia.htm> (last accessed: 10.04.2018).
- Pradilla, Miquel Àngel (2002), *La variació fònica en la llengua catalana: inventari i avaluació metodològica*, *Noves SL* (Autumn). <http://cultura.gencat.es/llengcat/noves/hemeroteca/metodologia.htm> (last accessed: 10.04.2018).
- Pradilla, Miquel Àngel (2003), *Les representacions del variacionisme en l'àmbit de la llengua catalana. Transferències i transaccions*, *Noves SL* (Autumn). <http://www6.gencat.net/llengcat/noves/hm03tardor.htm> (last accessed: 10.04.2018).
- Pradilla, Miquel Àngel (2008), *Sociolingüística de la variació i llengua catalana*, Barcelona, Institut d'Estudis Catalans.
- Pradilla, Miquel Àngel (2011), *L'univers calidoscòpic de la sociolingüística de la variació en la llengua catalana*, *Treballs de Sociolingüística Catalana* 21, 125-140.
- Pradilla, Miquel Àngel (2015), *La catalanofonia. Una comunitat del segle XXI a la recerca de la normalitat lingüística*, Barcelona, Institut d'Estudis Catalans.
- Pradilla, Miquel Àngel/Sorolla, Natxo (edd.) (2016), *Informe sobre la situació de la llengua catalana (2015)*. <http://www.demolinguistica.cat/arxiu/web/informe/informe2015.pdf> (last accessed: 10.04.2018).

- Romero, Sílvia (2001), *Canvi lingüístic en morfologia nominal a la Conca de Tremp*, Ph. D. thesis, Universitat de Barcelona.
- Suïls, Jordi, et al., (2010) *El canvi lingüístic al Pont de Suert, vint anys després (estudi en temps real a partir de l'estudi en temps aparent de 1986)*, in: *Actes del Quinzè Col·loqui Internacional de Llengua i Literatura Catalanes, Lleida, 2009*, Barcelona, Abadia de Montserrat, vol. 1, 297–314.
- Turell, Maria Teresa (ed.) (1995), *La sociolingüística de la variació*, Barcelona, Promociones y Publicaciones Universitarias.
- Turell, Maria Teresa (1997), *Variació i variacionisme*, in: Maria Rosa Lloret et al. (ed.), *Anàlisi de la variació lingüística*, Barcelona, Promociones y Publicaciones Universitarias, 45–64.
- Turell, Maria Teresa (2003), *El temps aparent i el temps real en estudis de variació i canvi lingüístic*, Noves SL (Autumn). http://www6.gencat.cat/llengcat/noves/hm03tardor/turell1_4.htm (last accessed: 10.04.2018).
- Verdú, Orland (2007a), *Canvi lingüístic en temps real: estudi del semiapitxament a Petrer*, Master dissertation, Universitat d'Alacant.
- Verdú, Orland (2007b), *Variation in real time: a case of sound change in Catalan*, presentation at the XVIII International Conference on Historical Linguistics, Montreal, Canada.
- Weinreich, Uriel/Labov, William/Herzog, Marvin I. (1968), *Empirical foundations for a theory of language change*, in: Winfred P. Lehmann/Yakov Malkiel (edd.), *Directions for historical linguistics: A symposium*, Austin, University of Texas, 96–195.
- Werlich, Egon (1975), *Typologie der Texte*, Heidelberg, Quelle & Meyer.

Joan Soler Bou

9 Language Corpora

Abstract: This chapter presents the main issues related to corpora in the Catalan language. It opens with a discussion of some general matters associated with corpora and empirical perspectives in linguistics, and provides a brief history of Catalan language corpora, taking into account the sociolinguistic conditions under which Catalan corpora have been developed which can explain the present situation and shortcomings in comparison with other languages. The main section provides a detailed inventory of the most important spoken and written corpora of the Catalan language, giving special attention to their structure, annotation and accessibility. Finally, the concluding remarks make some suggestions for how Catalan corpora should proceed in the future in order to maintain, or to regain, their rightful position among the rest of the European languages.

Keywords: corpus, corpus linguistics, Catalan corpora, spoken corpora, written corpora

1 Language corpora: the empirical perspective in linguistics

1.1 General issues

The goal of this paper is not to define what a corpus is, but to present the current situation of Catalan as far as language corpora are concerned. However, we should start such an undertaking by specifying what we consider a corpus to be. As most of the developments towards standardization in corpora compilation were made in the late 1990s, one option would be to look to the precise definition from one of the most relevant standardization teams in this period, the “Expert Advisory Group on Language Engineering Standards” (EAGLES). In the EAGLES documentation, Sinclair defines *corpus* and *computer corpus* as follows:

“A *corpus* is a collection of pieces of language that are selected and ordered according to explicit linguistic criteria in order to be used as a sample of the language. ... A *computer corpus* is a corpus which is encoded in a standardized and homogeneous way for open-ended retrieval tasks. Its constituent pieces of language are documented as to their origins and provenance” (Sinclair 1996).

Thus, following Sinclair, we do not consider textual repositories like collections and/or archives that do not fit this definition to be corpora. For a note on some of these discarded resources, see Massip-Bonet (2015) and Massip-Bonet/Llop-Naya (2015).

<https://doi.org/10.1515/9783110450408-016>

It should be pointed out that a very important feature of corpora is that they are supposed to be representative of a language or a certain subdomain of a language. The issue of representativeness (McEnery/Wilson 1996, 5) is probably the main source of discrepancy (Leech 1991, 8) between the concept of corpus posited by the post-Bloomfieldian linguists (such as Hill and Harris), criticized by Chomsky in the late 1950s, and the one established by the linguists who founded corpus linguistics a few years after. To say that any corpus must be representative of a certain state of things with regard to the language (or the languages) that it contains implies that it must fit the size, type of discourse, linguistic tagging, etc. required for a given research purpose. Despite the fact that some linguists, such as Léon (2010), consider the debate between introspective and empirical methods in linguistics to still be ongoing, the use of representative corpora is nowadays an undeniable necessity for the systematic description of any language, and also for the study of certain types of discourse.

In recent decades, the use of corpora in descriptive linguistics has increased. Corpora contribute to the study of language structure and use in most levels of linguistic analysis (McEnery/Wilson 1996; Biber/Conrad/Reppen 1998; Tognini-Bonelli 2001; Kübler/Zinsmeister 2015). Moreover, the evolution of electronic devices and computational techniques at the beginning of the 2000s made it possible to compile corpora containing larger amounts of linguistic data (hundreds of millions of words) than ever before. This trend shows no signs of having reached its limit. The huge number of texts being originally created in electronic form (that is, texts that do not have to be converted from paper to digital support), and the development of powerful tools to process and annotate linguistic information have made corpora of more than one billion words a possibility at the beginning of the 21st century.

As a result of these achievements, corpus linguistics has evolved (Tognini-Bonelli 2001) from a corpus-based approach (in which the corpus is used to expound, test or exemplify a pre-formulated linguistic theory) to a corpus-driven approach (in which any theoretical statement is provided by analysing the corpus itself, without predefined categories or patterns).

Finally, we should at least refer to the perspectives for corpus linguistics opened up by “Web as corpus” (Kilgarriff/Grefenstette 2003; Lew 2009), specially for the languages with a major presence on the Web.

1.2 The situation of Catalan

Our concern is to review how Catalan has created and maintained its own resources with regard to the evolution of linguistic analysis, and to evaluate which conditions limit and which facilitate this descriptive work as far as Catalan linguistics is concerned.

First of all, there are certain circumstances determining the situation of Catalan with regard to corpus linguistics that we should bear in mind. The first steps and the

achievements in the 1980s and the 1990s in terms of corpus creation, infrastructure and compilation of a reference corpus occurred not much later in Catalan than in other European languages. But these promising beginnings somehow lost their momentum in the 2000s, at least when it came to corpora being seen as research infrastructures. This was probably because the sociolinguistic situation (number of speakers, bilingualism, etc.) in the Catalan territories means that the only way of financing linguistic resources is through public research investment, meaning that there are no corpora promoted by publishing houses, etc. Nowadays, despite the relatively strong presence of Catalan online resources (electronic newspapers, web pages, Wikipedia, etc.), there are no large multifunctional corpora (running or in progress) in Catalan comparable in size, annotation and accessibility to those of the most influential languages in Europe.

As for the particularities of the current state of Catalan corpora, one should distinguish between institutional and research-oriented corpora. The former are linguistic resources, the aim of which is to provide data for the execution of internal collaborative work such as dictionaries, grammars, etc. Though their exploitation and use are primarily intensive and internal, these resources are usually publicly accessible through specific search interfaces, and they can be considered a research and social infrastructure. On the other hand, research-oriented corpora are compiled for more specific uses usually linked to university departments as a source of data for the execution of personal (PhD dissertations, etc.) and/or departmental research. In most cases, research-oriented corpora are accessible as raw or annotated data, without any search interface.

2 Catalan corpora: a brief history

2.1 Pre-electronic corpora

The history of corpora dates back to before the electronic era (Meyer 2008), when the most common goal of Catalan corpora was lexicography. Such corpora thus take the form of a repository of citation files with some comments, etc. Occasionally, these files have been published in dictionary form, as in the case of the *Diccionari Aguiló* (Fabra/de Montoliu 1914), which put the compilation of Marià Aguiló into dictionary form, or Miquel Colom's *Glossari general lul·lià* (Colom 1982–1985), which contains lexical information from the works of Ramon Llull.

In some cases, this type of repository has been converted into electronic form and is publicly available. This is the case for the *Nou glossari general lul·lià* (<http://nggl.ub.edu/glossari>), or the *Vocabulari de la llengua medieval de Lluís Faraudo de Saint-Germain* (<http://www.iec.cat/faraudo/>). In other cases, the corpus behind a lexicographic work may be published as a dictionary that may be converted into electronic form. Worthy of mention here, for instance, is the *Diccionari català-valencià-balear*

(Alcover/Moll 1926–1962) (<http://dcvb.iec.cat>), which is probably the Catalan dictionary based on the biggest pre-electronic corpus, containing 3 million citation files from about 1,100 texts.

Also worthy of mention, despite its methodological limitations, is the corpus compiled by Guiter (1972) listing the frequency of Catalan words (Rafel 1996). The list has 5,303 words obtained from a corpus of 50,494 words, taken from a small number of literary works, partially copied out.

2.2 Electronic corpora

The first attempts to create some infrastructure for developing computational resources from Catalan texts began in the late 1970s at the University of Barcelona (UB) (Badia 1977). The most important project was called “Prospecció automatitzada de textos catalans” (Rafel 1983). Thanks to the work done at the UB, the first dictionary and first basic components for corpus annotation were established. This project was also the start of the “Servei de Tractament Informatitzat de Textos Catalans”, which computed and processed a number of Catalan literary texts.

In 1983, the Institut d’Estudis Catalans (IEC) made the decision to launch a project called “Diccionari del Català Contemporani”, the first phase of which was the compilation of a reference corpus of contemporary Catalan texts (Rafel 1992) called *Corpus Textual Informatitzat de la Llengua Catalana* (CTILC). The experiences previously acquired in “Prospecció automatitzada de textos catalans” served as the basis for the project; for instance, the model of textual representation for CTILC texts (which were almost all typed manually) benefitted from the criteria and labels applied in the texts experimentally processed by the former project. Moreover, a first lexicon obtained from the inflectional expansion of the entries of the *Diccionari general de la llengua catalana* (increased later with the entries of the *Diccionari de la llengua catalana*, from Enciclopèdia Catalana) developed at the UB served as the core of the lemmatization system.

Following the start of the CTILC project, several other programs were launched in the late 1980s or the beginning of the 1990s. In general, these corpora were initiatives launched by universities, and for this reason their main aim was not institutional but research-oriented. In 1990 the Department of Catalan Philology at the UB launched the project *Corpus de Català Contemporani de la Universitat de Barcelona* (CUB). The CUB was formed of two divisions: on the one hand, a set of spoken corpora for various purposes (Boix 1996) and, on the other, a written corpus (Cabré/de Yzaguirre/Lorente 1996) called *Corpus Escrit de Català* (CECA) specifically designed for terminology extraction and studies on neologism.

The establishment, during these years, of research groups like the “Seminari de Filologia i Informàtica” (1986) at the Universitat Autònoma de Barcelona (UAB), centers like the “Institut Universitari de Lingüística Aplicada” (IULA) (1994) at the

Universitat Pompeu Fabra (UPF), and others that were integrated in the “Centre de Referència d’Enginyeria Lingüística” (1996–1998) represented a boost for Catalan’s linguistic resources. Many of these projects are presented in the following section along with other successful initiatives and projects related to corpus creation.

3 Existing corpora: classification and inventory

There are many criteria for classifying corpora (Torruella/Llisterra 1999; Rafel/Soler 2001) which we will not list in detail in this chapter, but an initial differentiation should be established upon the type of data contained in the corpus: spoken or written. Thus, spoken corpora are those composed of oral discourse (usually recorded) with a transcript that can be aligned in some way to the speech signal. As usual, we will distinguish between spoken corpora and speech corpora, which are collections of speech data used for developing, training or testing voice recognition/synthesis systems in linguistic engineering. We will not be including speech corpora in this paper; for a brief survey, see Bel/Marimon (2016).

Meanwhile, written corpora are made of texts that must be converted into the textual representation scheme of a particular corpus. Written corpora can usually be searched using interfaces that allow searches relating to linguistic units or phenomena across the whole corpus; taking into account that spoken corpora have transcriptions of texts, it is possible to browse such corpora in a similar way (see, for instance, Dahlmann/Adolphs 2009), but this is unusual, as we shall see in the case of Catalan spoken corpora.

3.1 Spoken corpora

Spoken corpora are usually smaller than written ones because they are more difficult to collect and process (Llisterra 1999). In the case of Catalan, transcriptions are often not aligned with sound, and cannot be accessed through an interface allowing global searches within the corpus.

3.1.1 The CUB (Corpus de Català Contemporani de la Universitat de Barcelona)

Under the acronym CUB (<http://www.ub.edu/ccub/>), the department of Catalan linguistics of the UB developed a series of spoken corpora (Boix 1996) of contemporary Catalan during the 1990s, each of them oriented to a particular linguistic/discourse descriptive area (dialectal, slang, formal/informal discourse, etc.). All these materials can be accessed and downloaded with a Creative Commons license from the service of digital resources of the UB (<http://diposit.ub.edu/>). This set of corpora has been used

in both personal (mainly PhD dissertations) and departmental research. The CUB was initially intended to integrate five corpora:

- a) *Corpus Oral Dialectal* (COD), Viaplana/Perea (2003): 49 short texts in recorded conversations, with phonetic and “phono-orthographic” transcriptions. Initially, 14 texts from the COD were published with alignment between transcripts and voice recordings in the CD-ROM of Viaplana/Perea (2003).
- b) *Corpus Oral de Conversa Col·loquial* (COC), Payrató/Alturo (2002): 10 short recorded conversations in colloquial language (about 15 min. each), transcribed in “discourse” form.
- c) *Corpus Oral de Registres* (COR), Alturo et al. (2004): 26 texts (of an average of 7 min.), transcribed in a discourse form, classified according to 7 different types of discourse: civil, educational-academic-scientific, political, private, commercial, cultural and religious.
- d) *Corpus de Varietats Socials* (COS), Boix et al. (2007): 25 texts from informants with the profile “manual worker” (except in one case labelled as “lumpen”). COS does not have, at the moment, any available materials.
- e) *Corpus Oral de Mitjans de Comunicació* (COM): it does not seem to have any available material or documentation.

Afterwards, the department of Catalan Philology presented as a complement to the CUB corpora the *Corpus Audiovisual Plurilingüe* (CAP-UB, Payrató/Fitó 2008). This is a multilingual corpus (Catalan, Spanish and English) made of recorded texts from 12 informants; texts are divided into 5 categories (narrative, descriptive, argumentative, expositive and instructive), and are obtained through two elicitation procedures (experiential and experimental). The texts (with transcripts and audio-visual support) were published in DVD format.

3.1.2 Other spoken corpora

There are four more spoken corpora:

- a) The *Clinical Interview Corpus* (CInt, Vila et al. 2010) is run by the Centre de Llenguatge i Computació (CLiC) at the UB. It is a bilingual Spanish-Catalan spoken corpus of clinical interviews aligned with the corresponding transcripts. The corpus consists of 15 hours of recordings divided into 40 clinical interviews (carried out in a hospital in Barcelona) lasting on average 22 minutes each. CInt is not publicly available but some information can be obtained from the CLiC webpage (<http://clic.ub.edu/corpus/en/clint>).
- b) The *Phonoprosodic Corpus of Spoken Catalan* (PhonCAT, Benet/Cortés/Lleó 2012) was developed in the Research Centre on Multilingualism at the University of Hamburg. It consists of 146 hours of recordings of read and elicited as well as spontaneous speech data from Catalan speakers in Barcelona. The speakers were

divided into nine groups according to their age and district. In general, only selected words were transcribed phonetically and orthographically, but certain recordings were fully transcribed orthographically. Access to the CInt files is restricted, and requires permission from the rights holder (<https://corpora.uni-hamburg.de/repository/spoken-corpus:phoncat>).

- c) *Glissando* (Garrido et al. 2013) is a bilingual corpus oriented to provide data on the prosody of Catalan and Spanish. It was developed by UPF, UAB and the University of Valladolid. Glissando consists of 25 hours of recorded speech from 8 speakers per language, and is comprised of two parts: read (news) and conversational texts. The entire corpus is transcribed orthographically, and annotated with the information relevant to prosodic studies, with a selective alignment. All the materials contained in the corpus are accessible under a Creative Commons License (<http://prado.uab.cat/glissando/>).
- d) Finally, the corpus *RETOC* (de Yzaguirre/Farriols/Martí 2007) is an initiative from the “Language Engineering Research Group” at the IULA-UPF. It is part of a larger research project linked to spoken corpus processing tools and comprises a total of 253 records with a total running time of 250 hours, of which 20 hours have been transcribed (according to the most recent data, from 2003). The materials were publicly available, but do not seem to be accessible currently.

3.2 Written corpora

3.2.1 Old-text corpora

Among the written corpora of old Catalan, a couple of projects are worthy of particular mention: the *Diccionari de Textos Catalans Antics* (DTCA) and the *Corpus Informatitzat del Català Antic* (CICA).

The DTCA is run by the “Centre de Documentació Ramon Llull”, which inherited the corpus resources from the now defunct “Servei de Tractament Informatitzat de Textos Catalans” mentioned above. DTCA is intended as the first stage of a more ambitious project called “Diccionari del Català Antic”. The corpus is comprised of 24 texts from the 13th to the 15th century, containing some 2 million words in total. The corpus is fully morphosyntactically (lemma and part of speech) annotated and manually validated. It is accessible through a well-documented search interface (<http://www.ub.edu/diccionari-dtca/>) that does not allow contextual filtering.

The CICA (Torruella 2009; Clavería/Torruella 2012) is run by the “Seminari de Lingüística i Informàtica” (UAB). It contains 414 texts from the 11th to the 15th century, containing a total of 8.6 million words. Texts are classified according to chronological, typological and dialectal criteria. CICA is accessible through a search interface (<http://www.cica.cat>) that allows contextual filtering. Though the corpus is apparently lemmatized by automatic means, the interface does not seem to make any

distinction between form and lemma. The CICA is the descriptive reference for the project “Gramàtica del català antic” carried out by Josep Martines and Manuel Pérez Saldanya.

Another corpus comprising old texts is the *Corpus Informatitzat del Valencià* (CIVAL), developed at the “Acadèmia Valenciana de la Llengua” (AVL). In fact, CIVAL combines old and modern texts, from the 13th to the 21st century (it is divided into two subcorpora: an old and a contemporary one). Although the size is not explicitly stated, the corpus is estimated to contain about 20 million words. Texts are classified according to chronological and typological criteria. CIVAL is morphosyntactically annotated (lemma and part of speech), apparently by automatic means, and is accessible through a poorly documented search interface (<http://www.ub.edu/diccionari-dtca/>) that does allow any degree of contextual filtering.

Finally, there are also two small corpora worth mentioning: the *Corpus Històric del Català* (HistoCat), developed by CLiC (Centre de Llenguatge i Computació) at the UB. There does not appear to be any documentation on the size, structure, etc., of the corpus but it seems to be quite small. It is automatically annotated, and accessible through a search interface (<http://clic.ub.edu/corpus/node/123>). The other is the *Corpus Dialectal del Català* (DialCat), which is similar to the former as far as creation, documentation, annotation, access, etc. is concerned. It appears to be comprised of pieces of orthographic transcripts from spoken dialectal texts.

3.2.2 Reference contemporary corpora

The most important corpus of contemporary Catalan is the *Corpus Textual Informatitzat de la Llengua Catalana* (CTILC) (Rafel 1994; 1996). It was developed between 1985 and 1997, as part of the programme “Diccionari del Català Contemporani”, at the Institut d’Estudis Catalans (IEC). The CTILC was conceived and executed under the direction of Joaquim Rafel and contains more than 52 million words, corresponding to 3,299 literary and non-literary texts dating from 1833 to 1988. In terms of linguistic information, the CTILC is fully morphosyntactically annotated (lemma and part of speech) by a semi-automatic procedure requiring manual validation. Moreover, in working with CTILC for lexicographic purposes, IEC experts have been able to detect and resolve tagging errors, which makes it a highly reliable reference corpus. It is accessible through a well-documented search interface (<http://ctilc.iec.cat>). The public interface allows lemma/form retrieval and chronological and typological filtering, but not contextual filtering. For research purposes, there is another interface that requires authentication, provided by the IEC on request, that allows the user to specify the context of the lemma/form being consulted.

As for internal institutional uses, CTILC has been the main source for the compilation of the *Diccionari descriptiu de la llengua catalana* (DDLC), completed in 2015 and fully accessible online (<http://dcc.iec.cat/ddlc/>). It is also used for various other

projects and purposes of the IEC (dictionaries and grammars). This major role of the CTILC, as well as its uniqueness as reference corpus of the Catalan language, have motivated the launch of a follow-up project, aiming at including texts from 1988 onwards in the CTILC. This continuation project entered its preparatory phase in 2015 and is expected to start its executive phase in 2018. Between 2015 and 2017, the IEC has been designing, implementing and adjusting the components necessary to the processing, tagging, storage and access to the expanded corpus. The project is expected to produce some visible and publicly accessible results in 2019.

During the compilation of the CTILC, the IEC also took part in the PAROLE project, devoted to the compilation of harmonized, generic and re-usable Written Language Resources for EU Languages. As part of this project, the IEC compiled the *PAROLE Catalan Corpus* (1996–1998), a 20-million-word corpus of general language texts (mainly journalistic texts and parts of the CTILC).

3.2.3 Language-specific corpora

There are at least three corpora composed of language-specific (or source-specific) texts that should also be mentioned:

- a) *Corpus tècnic de l'IULA* (Bach et al. 1997; Badia et al. 1998) is a multilingual corpus for language specific (terminology, detecting neologisms, etc.) purposes that contains specialised texts in the fields of law, economics, genomics, medicine, environment, plus a corpus of contrast media. The size of the Catalan subcorpus is 26.7 million words. The corpus is lemmatized by automatic means and is accessible through a search interface (<http://www.ub.edu/diccionari-dtca/>) that allows typological and contextual filtering.
- b) The *Wikicorpus-CAT* (Reese et al. 2010) has been compiled by TALP, the Research Center at the Universitat Politècnica de Catalunya (UPC). It is a trilingual corpus (Catalan, Spanish, English) taken from Wikipedia; the size of the Catalan subcorpus is 50 million words. The corpus is morphosyntactically annotated (lemma and part of speech) by automatic means, and also contains some semantic information (Word Sense Desambiguation techniques) using WordNet senses. It is not accessible through any user interface but is fully downloadable in tagged or raw text version (<http://nlp.lsi.upc.edu/wikicorpus/>).
- c) The corpus *El Català Escolar Escrit a Catalunya* (CesCa) was developed by CLiC (UB). It contains 2,426 texts produced by 5 to 9-year-old children and collected from 31 schools. The corpus contains texts belonging to four subject fields: movies, recommendations, jokes and definitions. The CesCa is automatically lemmatized and provides lists with information on the distribution of lemmas. It is accessible through a search interface (<http://clic.ub.edu/corpus/cesca-cerques-corpus>).

3.2.4 Syntactically and semantically annotated corpora

The corpora enriched with syntactic and semantic annotations are usually small, because of the level of human intervention they require. In Catalan there are two notable corpora of this type, which are both part of a bilingual corpus of Spanish and Catalan texts:

- a) *AnCorà-Ca* (Martí/Taulé/Recasens 2008) is a corpus of 500,000 words developed by CLiC-UB. It is based on journalistic texts, richly annotated at different linguistic levels (syntactic and semantic) and accessible through a search interface (<http://clic.ub.edu/corpus/ancora-cerques>).
- b) *Sensem Catalan corpus* (Castellón et al. 2006), developed by the “Grup de Recerca Interuniversitari en Aplicacions Lingüístiques” (GRIAL), was created by translating journalistic texts from the Spanish SenSem corpus into Catalan. It comprises approximately 600,000 words. It is accessible through a search interface (<http://grial.uab.es/sensem/corpus/main>) and is distributed on demand.

3.2.5 Parallel corpora

Parallel corpora are those composed of original texts and their respective translations, with both versions aligned according to segments or translation units (words, sentences, paragraphs, etc.). There are two main Catalan parallel corpora: the *Corpus Paral·lel Català-Castellà* based on the official journal of Catalan Government and Parliament, which includes 715,000 documents (from 1977 to 2015), and about 200 million words for each language. Meanwhile, the *Corpus Paral·lel Espanyol-Català* contains more than 100 million words (10 years of articles published in the newspaper *El Periódico de Catalunya*). The corpus is distributed by the “European Language Resources Association” (ELRA) with a license that allows for commercial use.

3.2.6 Non-accessible corpora

Among the corpora that are no longer accessible, it is worth mentioning the *Corpus d'Ús del Català a la Web* (CucWeb) (Boleda/Badia 2008). With a size of 166 million words, this was a corpus automatically compiled from the Internet, with a search interface allowing contextual filtering. Another corpus that is no longer accessible is the *Corpus Escrit de Català Contemporani* (CECA) (Cabré/de Yzaguirre/Lorente 1996) which was designed and executed as the written part of the aforementioned CUB.

3.2.7 Restricted corpora for specific purposes

The use of certain corpora is restricted to the internal purposes of their developers. Probably the most important and well documented is the *Corpus Informatitzat Multilingüe de Textos Antics i Contemporanis* (CIMTAC) (Martines/Sánchez 2014). This is an initiative of the Institute of Cooperative Research IVITRA, and comprises the following (sub)corpora: *Corpus Informatitzat de la Gramàtica del Català Antic* (CIGCA), *Corpus Informatitzat de la Gramàtica del Català Modern* (CIGCatMod), *Corpus Informatitzat Complementari del Català contemporani* (CICCatCo), *Corpus Informatitzat Multilingüe del Català* (CIMulCat), *Corpus Documentale Latinus Valencie* (CODOLVA).

4 Final remarks

The conclusion that can be drawn from the initial considerations and the inventory of corpora presented in this chapter is that Catalan is in a relatively strong position with regard to the linguistic resources that have been created so far. However, if we pay attention to the execution date of most of the corpora, some loss of momentum can be detected resulting from a lack of new projects.

There are clearly two aims, not exclusive but rather complementary, that Catalan corpora should be pursuing: firstly, to improve the existing corpora by enriching them with new annotations, increasing their accessibility, etc.; and secondly, to design and implement new, bigger and better corpora using a multifunctional approach that could be used in different kinds of linguistic and/or computational research and developments. If these goals are to be achieved, collaboration between institutions is needed, as well as the promotion of existing multidisciplinary teams and the creation of new research structures. This is the only way we will be able to overcome the limitations the Catalan language currently faces.

5 Bibliography

- Alcover, Antoni M./Moll, Francesc de B. (1926–1962), *Diccionari català-valencià-balear*, Palma, Moll, 10 vol., <http://dcvb.iec.cat> (last accessed: 22.09.2017).
- Alturo, Núria, et al. (edd.) (2004), *Corpus oral de registres. Materials de treball*, Barcelona, Universitat de Barcelona (Book and CD-ROM).
- AnCora-Ca = <http://clic.ub.edu/corpus/ancora-cerques> (last accessed: 22.09.2017).
- Bach, Carme, et al. (1997), *El corpus de l'IULA: descripció*, Barcelona, Universitat Pompeu Fabra.
- Badia, Antoni M. (1977), *Projecte d'informàtica lingüística de la Universitat de Barcelona*, in: Santiago Molfulleda (ed.), *Informàtica y lingüística. Mesa redonda sobre informática y lingüística (Barcelona, 26 y 27 de mayo de 1977)*, Barcelona, Fundesco, 7–21.
- Badia, Toni, et al. (1998), *IULA's LSP Multilingual Corpus: Compilation and Processing*, in: Nicoletta Calzolari et al. (edd.) *First International Conference on Language Resources and Evaluation*:

- Workshop proceedings, Granada 29–31 of May*, <http://www.iula.upf.es/corpus/corpubca.htm> (last accessed: 22.09.2017).
- Bake, Paul (ed.) (2009), *Contemporary Corpus Linguistics*, London/New York, Continuum.
- Bel, Núria/Marimon, Montserrat (2016), *Les indústries de la llengua i la tecnologia per al català*, *Llengua i ús*. Revista tècnica de Política Lingüística 58, 6–12, <http://www.gencat.cat/llengua/liu/58> (last accessed: 22.09.2017).
- Benet, Ariadna/Cortés, Susana/Lleó, Conxita (2012), *Phonoprosodic Corpus of Spoken Catalan (PhonCAT)*, in: Thomas Schmidt/Kai Wörner (edd.), *Multilingual Corpora and Multilingual Corpus Analysis*, Amsterdam/Philadelphia, Benjamins, 215–229.
- Biber, Douglas/Conrad, Susan/Reppen, Randi (1998), *Corpus Linguistics. Investigating Language Structure and Use*, Cambridge, Cambridge University Press.
- Boix, Emili (1996), *Els materials de llengua oral del corpus de català contemporani de la UB (CUB)*, in: Lluís Payrató et al. (edd.), *Corpus, corpora. Actes del 1r. i 2n. Col·loqui lingüístics de la Universitat de Barcelona (CLUB-1, CLUB-2)*, Barcelona, Universitat de Barcelona, 93–114.
- Boix, Emili, et al. (edd.) (2007), *Corpus de varietats socials. Materials de treball*, Barcelona, Universitat de Barcelona (book and CD-ROM).
- Boleda, Gemma/Badia, Toni (2008), *CUCWEB: un corpus de la llengua catalana construït a partir de la web2*, *Estudis Romànics* 30, 291–293.
- Cabré, M. Teresa/de Yzaguirre, Lluís/Lorente, Mercè (1996), *El projecte CECA (Corpus Escrit de Català)*, in: Lluís Payrató et al. (edd.), *Corpus, corpora. Actes del 1r. i 2n. Col·loqui lingüístics de la Universitat de Barcelona (CLUB-1, CLUB-2)*, Barcelona, Universitat de Barcelona, 115–125.
- Castellón, Irene, et al. (2006), *The Sensem Corpus: a Corpus Annotated at Syntactic and Semantic Level*, in: Nicoletta Calzolari et al. (edd.), *5th Edition of the International Conference on Language Resources and Evaluation (Genoa, 22–28 May 2006)*, <http://www.lrec-conf.org/proceedings/lrec2006> (last accessed: 22.09.2017).
- CesCa = *El Català Escolar Escrit a Catalunya*, <http://clic.ub.edu/corpus/cesca-cerques-corpus> (last accessed: 22.09.2017).
- Clavería, Gloria/Torruella, Joan (2012), *El “Corpus Informatizado del Catalán Antiguo” (CICA) y su herramienta de explotación (Estación de Análisis Documentales)*, in: María Jesús Torrens/Pedro Sánchez-Prieto (edd.), *Nuevas perspectivas para la edición y el estudio de documentos hispánicos antiguos*, Bern, Lang.
- Colom, Miquel (1982–1985), *Glossari general lul·lià*, 5 vol., Palma, Moll.
- CICA = *Corpus Informatitzat del Català Antic*, <http://www.cica.cat> (last accessed: 22.09.2017).
- CIVAL = *Corpus Informatitzat del Valencià*, <http://www.ub.edu/diccionari-dtca> (last accessed: 22.09.2017).
- CliC = *Clinical Interview Corpus*, <http://clic.ub.edu/corpus/en/clint> (last accessed: 22.09.2017).
- Corpus tècnic de l’IULA*, <http://www.ub.edu/diccionari-dtca> (last accessed: 22.09.2017).
- CTILC = *Corpus Textual Informatitzat de la Llengua Catalana*, <http://ctilc.iec.cat> (last accessed: 22.09.2017).
- CUB = *Corpus de Català Contemporani de la Universitat de Barcelona*, <http://www.ub.edu/ccub> (last accessed: 22.09.2017).
- Dahlmann, Irina/Adolphs, Svenja (2009), *Spoken Corpus Analysis: Multimodal Approaches to Language Description*, in: Paul Bake (ed.), *Contemporary Corpus Linguistics*, London/New York, Continuum, 125–139.
- DDLc = *Diccionari descriptiu de la llengua catalana*, <http://dcc.iec.cat/ddlc> (last accessed: 22.09.2017).
- de Yzaguirre, Lluís/Farriols, Antoni J./Martí, Jaume (2007), *El corpus RETOC: un corpus oral per a la recerca i la docència*, in: Sadurní Martí et al. (edd.), *Actes del Tretzè Col·loqui Internacional de Llengua i Literatura Catalanes*, Barcelona, Publicacions de l’Abadia de Montserrat, 495–504.

- DialCat = *Corpus Dialectal del Català*, <http://clic.ub.edu/corpus/node/123> (last accessed: 22.09.2017).
- DTCA = *Diccionari de Textos Catalans Antics*, <http://www.ub.edu/diccionari-dtca> (last accessed: 22.09.2017).
- Fabra, Pompeu/de Montoliu, Manuel (edd.) (1914), *Diccionari Aguiló. Materials lexicogràfics aplegats per Marian Aguiló i Fuster*, Barcelona, Institut d'Estudis Catalans, 8 vol., 1914–1934.
- Garrido, Juan María, et al. (2013), *Glissando: A Corpus for Multidisciplinary Prosodic Studies in Spanish and Catalan*, *Language Resources & Evaluation* 47, 945–971, <https://doi.org/10.1007/s10579-012-9213-0>.
- Glissando*, <http://prado.uab.cat/glissando> (last accessed: 22.09.2017).
- Guiter, Henri (1972), *Dictionnaire de fréquence du catalan*, *Via Domitia* 17, 13–49.
- HistoCat = *Corpus Històric del Català*, <http://clic.ub.edu/corpus/node/123> (last accessed: 22.09.2017).
- Iliescu, Maria/Roegiest, Eugene (edd.) (2015), *Manuel des anthologies, corpus et textes romans*, Berlin/Boston, De Gruyter.
- Kilgariff, Adam/Grefenstette, George (2003), *Introduction to the Special Issue on the Web as Corpus*, *Computational Linguistics* 29/3, 333–347.
- Kübler, Sandra/Zinsmeister, Heike (2015), *Corpus Linguistics and Linguistically Annotated Corpora*, London/New York, Bloomsbury.
- Leech, Geoffrey (1991), *The State of the Art in Corpus Linguistics*, in: Karin Aijmer/Bengt Altenberg (edd.), *English Corpus Linguistics*, London/New York, Longman, 8–9.
- Léon, Jacqueline (2010), *British Empiricism and Transformational Grammar. A Current Debate*, in: Douglas A. Kibbee (ed.), *Chomskyan (R)evolutions*, Amsterdam/Philadelphia, Benjamins, 423–444.
- Lew, Robert (2009), *The Web as Corpus Versus Traditional Corpora: Their Relative Utility for Linguists and Language Learners*, in: Paul Bake (ed.), *Contemporary Corpus Linguistics*, London/New York, Continuum, 289–300.
- Llisterri, Joaquim (1999), *Transcripción, etiquetado y codificación de corpus orales*, *Revista Española de Lingüística Aplicada*, Número Extraordinario 1, 53–82.
- Martí, M. Antònia/Taulé, Mariona/Recasens, Marta (2008), *AnCor: Multilevel Annotated Corpora for Catalan and Spanish*, in: Nicoletta Calzolari et al. (edd.), *6th Edition of the International Conference on Language Resources and Evaluation (Marrakech, 28–30 May 2008)*, <http://www.lrec-conf.org/proceedings/lrec2008> (last accessed: 22.09.2017).
- Martines, Vicent/Sánchez, Elena (2014), *L'ISIC-IVITRA i el metacorpus CIMTAC. Noves aportacions a la lingüística de corpus*, *Estudis Romànics* 36, 423–436.
- Massip-Bonet, Àngels (2015), *Catalan ancien: anthologies, corpus, textes*, in: Maria Iliescu/Eugene Roegiest (edd.), *Manuel des anthologies, corpus et textes romans*, Berlin/Boston, De Gruyter, 223–241.
- Massip-Bonet, Àngels/Llop-Naya, Ares (2015), *Corpus et anthologies du catalan contemporain*, in: Maria Iliescu/Eugene Roegiest (edd.), *Manuel des anthologies, corpus et textes romans*, Berlin/Boston, De Gruyter, 242–258.
- McEnery, Tony/Wilson, Andrew (1996), *Corpus Linguistics. An Introduction*, Edinburgh, Edinburgh University Press.
- Meyer, Charles F. (2008), *Pre-electronic corpora*, in: Anke Lüdeling/Merja Kytö (edd.), *Corpus Linguistics. An International Handbook*, vol. 1, Berlin/New York, De Gruyter, 1–14.
- Nou glossari general lul·lià*, <http://nggl.ub.edu/glossari> (last accessed: 22.09.2017).
- Payrató, Lluís, et al. (edd.) (1996), *Corpus, corpora. Actes del 1r. i 2n. Col·loqui lingüístics de la Universitat de Barcelona (CLUB-1, CLUB-2)*, Barcelona, Universitat de Barcelona.

- Payrató, Lluís/Alturo, Núria (edd.) (2002), *Corpus oral de conversa col·loquial. Materials de treball*, Barcelona, Universitat de Barcelona (Book and CD-ROM).
- Payrató, Lluís/Fitó, Jaume (edd.) (2008), *Corpus audiovisual plurilingüe*, Barcelona, Universitat de Barcelona.
- PhonCAT = *Phonoprosodic Corpus of Spoken Catalan*, <https://corpora.uni-hamburg.de/repository/spoken-corpus:phoncat> (last accessed: 22.09.2017).
- Rafel, Joaquim (1983), *Prospecció automatitzada de textos catalans*, Procesamiento del lenguaje natural 1, 24–32.
- Rafel, Joaquim (1992), *Cap a un diccionari del català contemporani*, in: Isidor Marí (ed.), *Segon Congrés Internacional de la Llengua Catalana. IV. Àrea 3: Lingüística social*, Palma, Universitat de les Illes Balears, 589–595.
- Rafel, Joaquim (1994), *Un corpus general de referència de la llengua catalana*, Caplletra 17, 219–250.
- Rafel, Joaquim (1996), *Introducció*, in: Joaquim Rafel (ed.), *Diccionari de freqüències 1. Llengua no literària*, Barcelona, Institut d'Estudis Catalans, I–CLIII.
- Rafel, Joaquim/Soler, Joan (2001), *El processament de corpus. La lingüística empírica*, in: M. Antònia Martí (ed.), *Les tecnologies del llenguatge*, Barcelona, Universitat Oberta de Catalunya, 27–59.
- Reese, Samuel, et al. (2010), *Wikicorpus: a Word-Sense Disambiguated Multilingual Wikipedia Corpus*, in: Nicoletta Calzolari et al. (edd.), *Proceedings of 7th Language Resources and Evaluation Conference*, La Valleta, <http://www.lrec-conf.org/proceedings/lrec2010/index.html> (last accessed: 22.09.2017).
- Sense Catalan corpus*, <http://grial.uab.es/sense/corpus/main> (last accessed: 22.09.2017).
- Sinclair, John (1996), *Preliminary Recommendations on Corpus Typology*, EAGLES Document EAG-TCWG-CTYP/P, <http://www.ilc.cnr.it/EAGLES96/corpus/corpus.html> (last accessed: 22.09.2017).
- Tognini-Bonelli, Elena (2001), *Corpus Linguistics at Work*, Amsterdam/Philadelphia, Benjamins.
- Torruella, Joan (2009), *Los ejes principales en el diseño de un corpus diacrónico: el caso del CICA*, in: Pascual Cantos/Aquilino Sánchez (edd.), *A Survey on Corpus-Based Research/Panorama de investigaciones basadas en corpus*, Asociación Española de Lingüística de Corpus, <http://www.um.es/lacell/aelinco/contenido/pdf/2.pdf> (last accessed: 22.09.2017).
- Torruella, Joan/Llisterri, Joaquim (1999), *Diseño de corpus textuales y orales*, in: José Manuel Bleca et al. (edd.), *Filología e informática. Nuevas tecnologías en los estudios lingüísticos*, Barcelona, Universitat Autònoma de Barcelona/Milenio, 45–77.
- Viaplana, Joaquim/Perea, M. Pilar (edd.) (2003), *Textos orals dialectals del català sincronitzats. Una selecció*, Barcelona, Universitat de Barcelona (Book and CD-ROM).
- Vocabulari de la llengua medieval de Lluís Faraudo de Saint-Germain*, <http://www.iec.cat/faraudo> (last accessed: 22.09.2017).
- Vila, Marta, et al. (2010), *CLInt: a Bilingual Spanish-Catalan Spoken Corpus of Clinical Interviews*, Procesamiento del Lenguaje Natural 45, 105–111.
- Wikicorpus-CAT*, <http://nlp.lsi.upc.edu/wikicorpus> (last accessed: 22.09.2017).
- Universitat de Girona, 9–12 de setembre de 2003*, vol. 1, Barcelona, Publicacions de l'Abadia de Montserrat, 495–506.

Language History

Philip D. Rasico

10 Early Medieval Catalan

Abstract: In this chapter the author examines in detail the formation and evolution of Catalan, from its Vulgar Latin origins to the beginning of the 13th century, based upon evidence provided by the earliest written manifestations of the language. Emphasis is given to phonological changes and morpho-syntactic developments in the transformation of spoken Latin to Catalan. Also considered are contacts with and the possible influence of pre-Roman languages upon the Latin introduced into the northeastern sector of the Iberian Peninsula in the late 3rd century BC, as well as the later impact of Germanic languages and of Arabic upon the development of emerging Catalan Romance. Evidence of the evolution of Early Medieval Catalan is provided by popular documents from the late 9th century to the beginning of the 13th century.

Keywords: origins, linguistic change, phonology, morpho-syntax

1 Historical antecedents

1.1 Foundational elements: The Pre-Roman and Roman periods

As in the case of its sister Romance languages, the origins of Catalan must be sought primarily, although not exclusively, in the particular variety or varieties of spoken, popular Latin introduced into the northeastern region of the Iberian Peninsula (also known as Hispania) in the late 3rd century B.C., together with the extreme southern portion of Gaul (mod. France) toward the end of the following century, by Roman legionnaires, public administrators, merchants, tradesmen, slaves and other members of the popular sectors of society.

It is well known that Rome's entry into and subsequent conquest and colonization of the peninsula occurred in response to the threat of territorial expansion north of the Ebro (Cat. Ebre) River by Carthage, its principal Mediterranean military and economic rival, through alliances established with various local indigenous groups. The arrival of Roman legionnaires in 218 B.C. at Emporion (Cat. Empúries), a small trading center established by Greeks in the 6th century B.C. on the Mediterranean coast just south of the Bay of Roses (Grk. Rhode), took place following the defeat of the Roman-allied Iberian city of Saguntum (Cat. Sagunt), located in the area of present northern Valencia, by forces led by Hannibal Barca, the Carthaginian military commander who, following the subjugation of Sagunt, proceeded northward through the Pyrenees and the Alps in a failed attempt to attack Rome (cf. Soldevila ²1963, 8–10).

<https://doi.org/10.1515/9783110450408-017>

In spite of much investigation by archaeologists, historians and linguists, little is known about the origins or the languages of the various groups or tribes that, prior to the Roman conquest, inhabited the area that would become Catalonia (Cat. *Catalunya*). In fact, the very origin of the terms *Catalunya* ('Catalonia') and *català* ('Catalan') remain unclear. Coromines (1954, 67–83; reproduced in Coromines 1965–1970, vol. 2, 159–174; and, with only minimal additions, also in *OnCat* vol. 3, 335–339, s.v. *Català, Catalunya*) appears to favor a corruption (metathesis) of the name of the pre-Roman Iberian tribe known as the *Lacetani* (or *Laketani*) that is thought to have inhabited extensive portions of central and eastern Catalonia. This tribe is mentioned as the *kastelanói* by the Greek geographer Ptolemy in the 2nd century A.D., a designation which, according to Coromines, is likely an error for **kattelanói*, the latter representing the source for the earliest known medieval references: *Cathalonia* in 1114, and *Catalania*, *Catalanicus*, *Catalaniensis* from approximately the same period, all of which appear in Italian sources. Coromines indicates, however, that there are also written attestations of related forms referring to several persons mentioned in documents from ca. 1112 that are included in the *Liber Feudorum Maior* (the manuscript of which is mainly from the 13th century): Raimundi *Catalan*, Arnal *Catalan*, and GERAL de *Cataluing*.

Nevertheless, it appears that when the Romans set foot upon the Iberian Peninsula at Empúries for the purpose of imposing Roman military and political control as well as Roman social institutions and along with the latter the Latin language, there were at least two and possibly three distinct ethno-linguistic groups within the future Catalan-speaking territory. These consisted of two distinct non-Indo-European-speaking groups along with a Celtic-based one: Iberian tribes, located along the eastern coast of the peninsula together with its immediate hinterland, whose territory extended northward from Murcia to approximately the lower Rhone River valley in southern France; Basque-related peoples dwelling north of the Montsec range and eastward along the Pyrenees from Ribagorça and Pallars through Andorra and Upper Urgell to the Cerdanya and perhaps beyond; and a blended Indo-European-speaking group known as Celtiberians, primarily of Celtic speech and culture but employing Iberian writing in their epigraphy, in the furthestmost western areas of the Catalan territory including the Ebro River valley and that of its tributary the Segre (Sanmartí 2009, 13–16; Lleal 1990, 18–28; Moran i Ocerinjauregui 1994, 13–14).

1.2 Romanization and Hispanic Latin

While the Romanization in the 2nd century B.C. of the territory that would become the *Conventus Tarraconensis*, an administrative division that included most of Catalonia, was intense, profound and accomplished rather rapidly (except in the Pyrenean valleys north of the Montsec together with Ribagorça), no less thorough was that of the southernmost area of the *Conventus Narbonensis* which included Roussillon (Cat.

Rosselló), albeit the latter was colonized by Rome nearly a century later (Badia i Margarit 1981b, 55–58). The complex process of Romanization, however, which initially was accomplished by means of the occupation of indigenous settlements by Roman legionnaires followed by the imposition upon the local population of Roman administrative, civil and cultural institutions including the Latin language, was not uniform. Urban nuclei located along the Mediterranean coast were those that were Romanized the most quickly and most efficiently and which subsequently served as centers for the spread of Roman rule and institutions. In addition to Emporiae (Empúries), these centers included Blandae (Blanes), Iluro (Mataró), Baetulo (Badalona), Dertosa (Tortosa), Ruscino (Castell-Rosselló), and especially Tarraco (Tarragona) and Barcino (Barcelona). In the interior of Catalonia only Ilerda (Lleida) and perhaps also Guissona seem to have played a similar role, although to a far lesser extent. However, a third sector of Catalonia, situated north of the Montsec mountain range and west of the Segre River and comprising the geographically isolated regions of Ribagorça and Pallars as well as the Vall d'Aran and Andorra, appears never to have experienced Romanization, since in these areas a Basque-like language survived until several centuries after Romance (primarily Catalan; but Occitan in the case of the Vall d'Aran and Aragonese in the western part of Ribagorça) already had emerged from the Latin spoken in neighboring regions (Soldevila 1963, 15–16; Badia i Margarit 1981b, 62–68). As Russell-Gebbett (1965, 12) notes:

“[I]t seems probable that the Basque-speaking tribes of the north-west were but little affected linguistically by their inclusion in the juridical convent of Tarragona: they stood apart from the main routes linking north-east Tarraconensis with Gallia Narbonensis (and with Italy), and acquired their ‘Latin’ considerably later than the regions to their east.”

In contrast, elsewhere in Catalonia pre-Roman languages seem to have disappeared quickly in major urban centers along the coast, while in the more remote rural villages of the interior they apparently survived until at least the 1st century A.D., albeit undoubtedly in an increasingly diglossic situation alongside Latin. According to Arasa i Gil (1994–1995, 87–88):

‘The death of the Iberian language may be explained as a case of the replacement of a local language by another of a more efficient nature and of a broader context [...]. Prior to the adoption of Latin by the indigenous populations, there occurred a period in which the latter preserved their language for colloquial use and employed Latin for special occasions, in a bilingual situation – strictly diglossal – in which the local languages occupied a position of decreasing prestige leading to their disappearance’ (translation PDR).

The Latin introduced through Roman conquest into the extreme northeastern sector of the *Hispania Tarraconensis* and the southeastern portion of the *Gallia Narbonensis* (a region later known as *Septimania*), a language which was to serve as a common means of communication among the various ethno-linguistic populations under the control of Rome, must have been characterized by considerable diversity in spite of

the relative uniformity of the written language. Little is known about the origins of the Romans who settled in the aforementioned regions and whose Latin undoubtedly presented significant geographical, socio-cultural and chronological differences according to the speakers' provenance, ethno-linguistic background and social status, as well as the time period in which they arrived. In this regard Lleal (1990, 36–46) observes that among the complex factors responsible for the diversification of spoken Latin in the various geographical regions of the Roman Empire, the following are especially noteworthy: a) the period of Romanization; b) the ease of communications with Rome; c) the degree of influence of indigenous languages (i.e. the linguistic substrate); d) the sociolinguistic features of those responsible for transmitting the language (i.e. diatopic and diastratic variation); and e) the influence of other languages upon Latin following Romanization (i.e. the linguistic superstrate).

It is clear that in spite of the difference in the date, but not in the degree, of their Romanization, the Latin spoken on both sides of the Eastern Pyrenees, the areas in which popular Latin would evolve into Catalan as a constitutive Romance language, shared a common linguistic base a number of features of which, especially of a lexical and phonological nature, were also proper to varieties of Latin introduced into more northern regions of Septimania and beyond (Veny ⁹1991, 19–20; Russell-Gebbett 1965, 12). Although they have been and are still the subject of debate among scholars, among those factors that may have contributed to the trans-Pyrenean linguistic ties mentioned are a) a common pre-Roman substrate, and b) a possible Latin base whose speakers originated in areas of Italy where Italic languages different from Latin, such as Oscan and Umbrian, had been spoken prior to the incorporation of those regions into the Roman Empire (Bastardas 1995, 49, 60–64).

1.3 Germanic Invasions and dissolution of the Roman Empire

Beginning in the 3rd century A.D. the frontiers of the Roman Empire faced the threat of incursion by various barbarian peoples, a situation which, coupled with serious internal problems within the Empire, set in motion a series of dramatic transformations of a political, socio-economic and cultural nature that led not only to Diocletian's division of the Empire into two distinct geopolitical entities in the late 3rd century, but also eventually to the further fragmentation of centers of power and culture and ultimately to the collapse of the Empire in the late 5th century. During this period groups of Germanic peoples (Franks, Alamanni, Suebi) originating in the Baltic Sea and Elbe River regions began crossing the frontiers of the Roman Empire established along the Rhine and Danube Rivers. Many of these peoples were initially absorbed into the late Empire as mercenaries posted along the frontiers or as colonists established in those same areas. The 4th and 5th centuries witnessed mass incursions of additional Germanic peoples from the Baltic region (Goths, Vandals, Alans) into the territories of the imploding Empire.

Among the Germanic groups who migrated en masse to the territories of the flagging, decentralized Empire in the 4th century were the Visigoths ('Western Goths') whose origins are thought to have been in Scandinavia. Pushed westward by the Huns, the Visigoths initially were welcomed by the Romans as confederates of the late Western Roman Empire; and, following periods of conflict with Rome including the sacking of the capital of the Empire in 410, they ultimately underwent a gradual process of Romanization that included the official adoption of Latin along with that of the Christian faith in its Arian form. Early in the 5th century, in an effort to maintain control of Hispania, Rome made an agreement with the Visigoths whereby, in exchange for their continued collaboration in the defense of the Western Empire against the invading Vandals and Alans, they were granted rule over Aquitaine and western Languedoc in southern Gaul, and where later in that same century they established the Visigothic kingdom of Toulouse whose territory came to include Septimania and much of Hispania. However, faced with Frankish expansion from the north in the early 6th century, the Visigothic kingdom in Gaul was forcibly restricted to Septimania (the Narbonensis) and, south of the Pyrenees, it was initially limited to the northeastern sector of the Tarraconsensis from which it spread to Toledo (the capital of the Hispano-Visigothic kingdom) and to much of the rest of Hispania in the latter part of the same century (Soldevila 20163, 2–27; Lleal 1990, 58–63).

The collapse of the Western Roman Empire and the establishment of a Romanized (i.e. Latin-speaking) and, apparently, initially bilingual Germanic people on both sides of the Eastern Pyrenees were to have significant consequences for the evolution of the locally spoken Latin to Catalan, and especially in the early formative period of this trans-Pyrenean Romance language corresponding approximately to the 6th through the 8th centuries. According to Russell-Gebbett (1965, 12–13):

"In the fifth century the Visigothic 'allies' of Rome confirmed the frontiers of the Tarragonese juridical convent, but renewed the bonds between its north-eastern area and Narbonensis [...]. [A]t the beginning of the formative period of the Catalan language, Septimania was united politically with Tarraconensis; an important span in the bridge between this province and Narbonensis had been rebuilt, and the cultural traffic across it was to run henceforth for many centuries in a predominantly north-south direction. A period of strong Provençal influence upon nascent Catalan had begun."

Bastardas (1995, 64, 71), who argues in favor of a substrate-based explanation for the trans-Pyrenean origin of Catalan, observes the following regarding the aforementioned links between the territories lying both north and south of the Eastern Pyrenees:

"It seems to me that, thanks to the common substrate, and with this designation I refer not only to an ancient linguistic community but also an ethnic and cultural one that spoke the same Romance language on both sides of the extreme Eastern Pyrenees, in spite of chronological differences regarding their Latinization [...]. It is perhaps not a trivial fact that the area of influence of Greek Empúries included all of Roussillon. Consequently, when there is an ancient

and deeply established community, a constant process occurs of linguistic levelling [...] An uninterrupted relationship with the Occitan lands, during the Roman period as well as the Visigothic, explains the parallel development of the Catalan and Occitan territories. This relationship only deepened when the Catalan counties were incorporated into the Carolingian Empire; however, from the very first written examples the two languages appear clearly distinct. A few characteristic features of Catalan may be due to the Italic substrates introduced by the first colonizers. However, the role of the pre-Roman substrates must have been very important, especially in the case of the Basque-like languages. This effect of the substrates must have contributed to the particular structure of Catalan' (translation PDR).

1.4 Germanic influence upon emerging Catalan Romance

The linguistic impact of the Visigoths upon the developing Romance language of Catalonia appears to have been minimal and primarily lexical. Many of the words of Germanic origin in Catalan and other Romance languages of Hispania were in fact already introduced into spoken Latin during the Roman period (e.g. WERRA > *guerra* 'war', *BLANK > *blanc* 'white', *RIKS > *ric* 'powerful, rich', *FRISK > *fresc* 'fresh', *WAIDANJAN > *guanyar* 'to gain', *SUPPA > *sopa* 'soup'), while numerous additional lexical items were incorporated into early Catalan Romance, especially from Frankish and also through Occitan (e.g. FRANK > *franc* 'free', *WIDAN > *guiar* 'to guide', *STAKKA > *estaca* 'stake', BUSK > *bosc* 'forest', KNIF > *ganivet* 'knife', *BLATU > *blat* 'wheat', *KROSTJAN > *cruixir* 'to crunch', *WANTA > *guant* 'glove', BLAO > *blau* 'blue'; or the productive suffix -ING > -enc 'of' indicating notions such as relationship or origin) (Lleal 1990, 100–101; Badia i Margarit 1981a, 24–25; Moll 1991, 59). However, in addition to the latter and unique morphological element, it is possible that the phonological system of Early Catalan, and especially in relation to the development of stressed and unstressed vowels in the eastern territories of Old Catalonia, was influenced by the Germanic superstrates of spoken Latin, that is by Visigothic as well as Frankish, since it is known that the languages of these Germanic groups possessed an especially notable system of heavy stress that may have led to more frequent instances of atonic vowel loss (vowel syncope as well as apocope) in languages such as Catalan, Occitan and French, as opposed to Castilian, Asturian-Leonese, Galician-Portuguese, etc. (Rasico 1982, 115; 2006, 143–145, 257; Moran i Ocerinjauregui 1994, 1–20).

Although the contribution of Germanic to spoken Catalan was relatively modest overall, in the period embracing the 9th through the 11th centuries a considerable portion of the population of Old Catalonia, that is of the northeastern regions including Roussillon, appears to have borne personal names of Germanic origin (primarily Visigothic) many of which continue today as surnames, such as *Arnal/Arnau* < ARNOALD, *Bertran* < BERTRAND, *Galí* < GALINDU, *Gombau* < GUNDOBALD, *Gubert* < GOADPERT, *Llopart/Llompert* < LEUPARD, *Requesens* < RICOSIND, *Tubau/Tibau* < THEUDIBALD, etc. (Badia i Margarit 1981a, 25–26; Moll 1991, 59–60). Moreover the toponymy of Old Catalonia, and especially that of its northeastern sector, reveals an

abundance of place names whose origins may be traced to Germanic personal names or to hybrid formations consisting of a Romance term derived from Latin followed by a Germanic name, as in *Arderiu* < HARDARĪCI, *Gisclareny* < GISCLASIND, *Guitarriu* < WITIRĪCI, *Gombrèn/Gombreny* < GUMESIND, *Escariu/ Escriu* < ASCARĪCI, *Ralleu* < ADALHAID, (*la*) *Geltrú* < WISALTRUD; *Castell de l'Areny* < CASTELLU ATHALASINDI, *Vilopriu* < VĪLLA HILPERĪCI, *Montsoriu* < MONTE SIGERĪCI (Badia i Margarit 1981a, 26; Moll 1991, 60; Coromines 1965–1970, vol. 1, 26, 31–65). In contrast to the northeast, in regions of western Catalonia such as les Garrigues, la Noguera, el Pla d'Urgell, la Segarra, el Segrià and Urgell, the incidence of toponyms of Germanic origin is appreciably inferior to that of the former area (Turull 2007, 308); and in the Valencian territory the few toponyms traceable to Germanic etyma appear to have originated in Old Catalonia from whence they were transferred to the south following the reconquest of this extensive area in the mid-13th century (Badia i Margarit 1981a, 26; Moll 1991, 60).

1.5 Transformative events of the Early Middle Ages

The numerous civil conflicts that characterized Hispania from the mid-7th to the early 8th centuries, the consequences of which led to the fracturing of the relative political unity established by the Visigothic monarchy in the 6th century, ultimately opened the doors of the peninsula to incursions by militant forces consisting of Berbers and Arabs from northern Africa who, as of 711, quickly occupied three quarters of the Iberian Peninsula and put an end to Visigothic rule. These groups introduced into Hispania not only a radically distinct ethnicity, religion and culture, but also a new non-Indo-European language, Arabic, that would have a major impact upon the Catalan territories located south of the Llobregat River and especially in the Segre and Ebro basin areas.

The northernmost regions of Old Catalonia and Septimania, however, experienced only a short period of Arab control. Under Frankish leadership Narbonne along with the rest of Septimania was wrested from Arab control in 759. And only a few decades later Frankish forces similarly began the conquest of the Catalan territories immediately south of the Pyrenees: Girona was freed in 785; Vic, Cardona and Casserres were repopulated in 798; and in 801 the important port city of Barcelona fell to the Franks under the leadership of Charlemagne's son, Louis the Pious (Soldevila ²1963, 37–40; Bisson 1988, 29–31).

Concerning the reconquest and repopulation of the northern areas of Old Catalonia, Soldevila (²1963, 37) notes:

'Septimania was the region of refuge for those Christians from our lands who refused to remain under the Saracen yoke. They settled principally in the Roussillon and in the areas of Narbonne, Carcassonne and Béziers [...]. Refugees from Septimania encouraged and assisted in the Reconquest, and when they returned to Catalonia (those who managed to return), after having lived

among the locals, they took with them elements of repopulation. That period of people from our land residing in Septimania represents a new connection whose importance, from multiple perspectives, including linguistic consideration, perhaps has not been sufficiently taken into account' (translation PDR).

Regarding the Frankish-led reconquest of northern Old Catalonia, Russell-Gebbett (1965, 13–14) observes:

"Catalonia north of the Llobregat (*Catalunya Vella*) had by the dawn of the ninth century been reconquered and resettled, and divided for administrative purposes into counties: Girona, Empúries, Barcelona, Ausona, Urgell and Cerdanya; these formed with Septimania the Marquisate of Gotia, while Pallars and Ribagorça – never conquered by the Moors – were incorporated into the Marquisate of Toulouse. In 864 the Catalan counties of Gotia were detached from Septimania to form a separate entity, the Marca Hispanica, Pallars and Ribagorça remaining for another twenty years under the mantle of Toulouse."

As a consequence of the brief period of Arab control over much of northern Old Catalonia, the linguistic impact of Arabic upon the local toponymy as well as the emerging Catalan language was either nonexistent or minimal. In general it may be said that among the Romance languages of the Iberian Peninsula, the impact of Arabic upon Catalan was the least notable. In fact, many if not most of the Arabic contributions to the Catalan lexicon appear to have originated in the southernmost regions of the Catalan linguistic territory, and especially in the Ebro River basin and, as previously noted, also in the Segre River valley (Badia i Margarit 1981a, 26–27; Moll 1991, 60–62; Turull 2007, 308–309). Examples of Arabic terms incorporated into the Catalan lexicon cover a wide range of semantic categories including plants, fruits and grains (*albergínia* 'eggplant', *alfàbrega* 'basil', *arròs* 'rice', *atzavara* 'agave', *carbassa* 'pumpkin', *carxofa* 'artichoke', *cotó* 'cotton', *garrofa* 'carob', *llimona* 'lemon', *taronja* 'orange'); agriculture (*aljub* 'cistern', *sèquia* ~ *síquia* 'irrigation canal', *sènia* ~ *sínia* 'waterwheel', *safareig* 'small reservoir', *sucre* 'sugar'); hydrographic accidents (*albufera* 'saltwater lagoon', *rambla* 'stream', etc.); names of winds (*garbí* 'southwest wind', *xaloc* 'southeast wind'); construction (*alcova* 'alcove', 'bedroom', *golfà* 'attic', *ràfec* 'eaves', *rajola* 'tile'); kitchen-related terms (*aixeta* 'tap, faucet', *gerra* 'jar', 'pitcher', *safata* 'tray', *tassa* 'cup'), among numerous additional fields (Badia i Margarit 1981a, 26–28; Moll 1991, 60–62).

It appears that the early counts of the reconquered territories in Septimania and Old Catalonia were appointed by the Frankish king. One of these early counts, Sunifred of Carcassone, a descendant of Visigoths, led the campaigns that successfully freed the regions of Cerdanya (835) and Urgell (838). His oldest son, Guifré el Pilós (870–897), who was given jurisdiction over Urgell, Cerdanya and Conflent in 870, and later also over Barcelona and Girona in 878, is considered the last of the Catalan counts of the Marca Hispanica to have been appointed by royal decree. Upon his death in 897 while resisting a Moorish assault on Barcelona, Guifré's sons divided up his counties among themselves: the eldest, Guifré II 897–911, took over control of

the counties of Barcelona, Girona and Osona; Miró II (897–927) assumed administration of Cerdanya, Conflent and Berguedà; while Sunifred II (897–950) became the count of Urgell. Upon the untimely death of Guifré II his youngest brother, Sunyer (911–947), received Barcelona in a highly unusual move that apparently resulted in Miró II receiving Besalú in 913. Although the Catalan counts would continue to acknowledge and pledge loyalty to the Frankish king – whose protection they occasionally found necessary – until the late 10th century, by the end of this period they were clearly governing their respective domains independently (Soldevila 1963, 59–86; Bisson 1988, 29–33).

Among Frankish contributions to the future territory of Catalonia was the implantation of a feudal system which introduced a series of socio-economic and political changes that would have a decisive impact upon the future development of this territory and its people. The Mozarabic liturgical rite was replaced by the Frankish, and Catalan monastic institutions reestablished close ties with their trans-Pyrenean counterparts. Moreover, in the *scriptoria* of those same monasteries the Visigothic script was replaced by the Frankish; Catalan documents were dated by the reign of Frankish kings until 1180; and the circulation of Carolingian-minted coins was common (Russell-Gebbett 1965, 15). Soldevila (1963, 45–46) summarizes the influence of the Franks upon Catalonia as follows:

‘The Frankish Reconquest had only the slightest impact from an ethnic perspective, since the Franks were principally among the ruling classes. However, in linguistic terms, while it left only a limited number of words in our language, it proved transformational in other respects. Because of it, Catalonia felt little in common with the rest of the Peninsula. The struggle with the Saracens, in the first centuries of the Reconquest, far from separating it from the lands of southern Gaul, strengthened its ties with the latter; far from isolating it from Europe, it placed it in more direct contact; far from tying it to the rest of Spain, it loosened its connections; to the point that, for other peninsular peoples, those of our land would be known, for quite some time, as *the Franks*’ (translation PDR).

The Catalan reconquest of Moorish-held territories moved slowly to the south while, to the west and southwest it progressed much more rapidly especially under Ramon Berenguer I of Barcelona (1035–1076) who, allied initially with Ermengol III of Urgell (1038–1066), extended Catalan dominion westward in a successful effort to limit the power and influence of Moorish-held Lleida while, at the same time, blocking the southern expansion of Pallars and Urgell as well as a possible eastward spread of the Kingdom of Aragon. Camarasa and Àger (Noguera) were reconquered in 1050; Pilzà, Purroi and Estopanyà (Lower Ribagorça) were captured in 1063; and Cervera (Segarra) was taken in 1070. However, Tarragona would not fall decisively under Catalan control until 1118; and it would take several more decades before Lleida surrendered and was repopulated in 1149. By the mid-12th century the Catalan political and linguistic territory had come to include the important city of Tortosa (1148) together with Fraga (1149) and most of the Ebro River valley (Soldevila 1963, 99–104, 147–169; Russell-Gebbett 1965, 16–17).

With the reconquest and occupation of these former Moorish towns and regions the Catalan language quickly eliminated and replaced what vestiges may have remained of former local Romance vernaculars, collectively known as Mozarabic or *Andalusí* Romance, some evidence of which may be still found in both toponyms and personal names appearing in documents corresponding to this period (Sanchis Guarner 1950, 99–144; 1961, 103–150; 1972, 118–123; Gili Gaya 1955; Ferrando Francés/Nicolás Amorós 2011, 75–82).

2 Appearance of Catalan in written sources

There is no doubt that when the Franks began the Reconquest of the Catalan territories located south of the Eastern Pyrenees in the late 8th century, the inhabitants of these regions were already speaking an early and evolving form of Catalan Romance. The earliest written manifestations of the Catalan language, however, are not discovered until the 9th century in late-Latin manuscripts and consist primarily of personal names, and especially of place names, whose Latin etyma (if in fact they were of Latin origin) appear to have been less than transparent to those scribes who recorded the documents and whose knowledge of Latin must have been only superficial or at best limited to common phrases and formulas.

Bastardas (1995, 91, 110–111) offers the following observations concerning the evolution of local spoken Latin to Catalan:

'More interesting than determining at which point that which is being spoken may no longer be considered Latin is, perhaps, determining when the conscious realization occurs that what is spoken constitutes a language different from Latin; the appearance of this consciousness (naturally by those individuals who had some knowledge of the Latin language) must be placed at the end of the 8th century rather than at the beginning of the 7th [...]. It was precisely the Carolingian reform, which made Latin an academic language subject once again to strict norms, that gave rise to the consciousness that Latin and the spoken language were two very different and distinct linguistic entities [...]. The language spoken by the Barcelonans and Gironans during the Saracen occupation certainly was not Latin, nor was Latin the colloquial language of the Hispanic exiles in Septimania, following the Muslim repression of those who had collaborated in Charlemagne's failed expedition to Saragossa in 778, nor of the Gironans who in 785 voluntarily delivered their region to Charlemagne, nor of the inhabitants of the regions of Cerdanya and Urgell who did likewise a short time later. The Barcelonans freed in 801 spoke a Romance language, and this Romance language would never cease to be just that. Although exposed to external influences and pressures, it would mainly evolve according to its own dynamic' (translation PDR).

Sources from the 9th to the 13th centuries that provide written evidence of emerging Catalan include: consecrations of churches; inventories of various types (e.g. lists of property belonging to episcopal sees or to individual churches, monasteries or convents; *capbreus* or lists of fiscal obligations, etc.); donations; wills and testaments; acknowledgements of property rights and receipts for the sale or transfer of property; legal agreements including feudal oaths and treaties; and especially documents

expressing grievances (*querimònies* ‘accusations’, *clams* ‘complaints’ and *rancures* ‘resentments’) related to the transgression of feudal rights and customs or to other forms of alleged abuse committed by feudal lords against their vassals or vice versa (Moran i Ocerinjauregui 1994, 55–93; Rasico 2015). With regard to the use of Catalan as of the 11th century in documents relating to feudal society, Moran/Rabella (2015, 121–122) observe:

‘[T]he consolidation of feudalism is the main reason for the Romance language’s access to the written code. Although until that point written language was a medium reserved for Latin, the formation of new events, values and concepts that no longer formed part of the Latin world and which, therefore, did not have a fixed name in that language, made these new elements filter into the written language [...]. This process was also facilitated by the low level of culture of the period of the beginning of feudalism, with little general knowledge of Latin outside the most enlightened ecclesiastical centers. The fact was that, in a society which, contrary to what we might think nowadays, generated a great deal of documentation, scribes, faced with the difficulty of translating these new situations and the statements in the mouths of lay people into Latin, had no option when drawing up their documents but to give the basic reasoning directly in Romance’ (translation PDR).

While Latin continues to be the predominant written language throughout the 12th century, the use of Catalan, or that of a thinly disguised (i.e. Latinized) vernacular, becomes increasingly more common throughout this century and the early decades of the following one, a period that would witness the translation into Catalan of legal documents and religious texts such as the *Liber Iudiciorum* (12th cent.) and the *Homilies d’Organyà* (early 13th cent.) (Moran i Ocerinjauregui 2006, 43–48; Moran/Rabella 2015, 119–121).

3 Outline of preliterate Catalan phonology

3.1 Development of vowels

3.1.1 Stressed vowels

Among the earliest Catalan documents that have been preserved to date is an original manuscript that purports to record the official act of consecration of the restored cathedral of Santa Maria at la Seu d’Urgell in 819 or 839, but which in fact is a falsification dating from the second half of the 9th century.¹ This document, which

¹ Baraut (1984–1985, 524–525) estimates the date of composition of the manuscript to be ca. 860–880. See also Moran i Ocerinjauregui (1994, 55–56). Russell-Gebbett (1965, 223), who follows Pujol (1917, 95–96), notes: ‘The restored cathedral of Urgell was endowed in 839, the Act of Consecration listing 289 localities subject to its jurisdiction: 129 in Alt Urgell and the Solsonès, 85 in Cerdanya and the Vall de Ribes (upper Ripollès), 31 in the Berguedà, 42 in Pallars and 2 in Ribagorça.’

identifies 289 villages or parishes under the jurisdiction of la Seu d'Urgell, contains numerous toponyms whose forms clearly reveal phonological developments proper to Catalan, such as the simplification of the stressed Vulgar Latin or Romance diphthong AI to *e* (/e/) via the stages [aj] > [ej] > [e] > [e]. Examples of this result in the aforementioned text are *Ferrera* < FERRARIA, *Napinerios* < NAPINARIOS, *Banieras* < BALNEARIAS, *Argilers* < ARGILLARIOS, *Siskero* < *SISKARIU, *Terreros* < TERRARIOS, *Traverseras* < TRANSVERSARIAS, *Tartera* < TARTAREA, *Kerosalbos* < *CARIOS-ALBOS, *Kaballera* < CABALLARIA, *Macianeros* < *MATTIANARIOS, *Argilageros* < *AIELAG (A)-ARIOS, *Kerubio* < *CARIU-RUBEU, *Palomera* < PALUMBARIA (cf. *Palomerola* < PALUMBARIOLA), *Ribera* < RĪPARIA (Baraut 1978, 51–52). Numerous additional examples of the monophthongization of AI to *e* from similar Latin sources (-ARI-, -ARE-) appear in Catalan texts as of approximately the same period: *Corbera* 899 < CORVARIA (Baraut 1978, 68), *Viadero* ca.950 < VIADARIU (Mundó 1974, 239), *aperos* 960 (Baraut 1978, 95)/*apers* 986 < *APPARIOS (Russell-Gebbett 1965, 62), *Colomera* ca. 967 < COLUMBARIA (Rasico 2006, 101), *Brugera* ca. 997 < *BRŪCARIA (Rasico 2006, 101–102), *cavalers* 1036–1079 < CABALLARIOS (Russell-Gebbett 1965, 65), *quartera/quartères* ca. 1075 < QUARTARIA(S) (Rasico 2006, 305), *Molner* ca. 1184 < MOLĪNARIU (Rasico 2006, 324).

The reduction of Vulgar Latin or early Old Catalan *ai* ([aj]) to *e* ([e]) from other sources is also found: *fexa* ca. 950 < FASCIA (Mundó 1974, 239); *dret* 12th cent. < DIRĒCTU, *esplets* 12th cent. < EXPLIC(I)TOS (Rasico 1993, 115); *seré* 1028–1047 < *ESSERE-HAIO (Russell-Gebbett 1965, 63), *faré* 1035–1076 < *FACERE-HAIO (Rasico 2006, 102–103); *comanné* 1080–1095 < COMMANDĀVĪ, *doné* id. < DONĀVĪ, *rancuré* id. < *RANCŪRĀVĪ (Russell-Gebbett 1965, 74–75). Nevertheless, the simplification of *ai* to *e* does not appear to have been uniform either chronologically or geographically, as in the case of -ACT- > *eit* > *et* ([akt] > [ajt] > [ejt] > [ejt] > [et]) in which the diphthong is still retained in the speech of several northwestern Catalan regions and especially in that of Pallars and Ribagorça (Coromines 1976–1977, vol. 1, 101–108; Veny 1991, 131). Some examples of the conservation of AI > *ei* ([ej] or [ej]) are *Corbeira* ca. 995 < CORVARIA (Rasico 2006, 102), *Cabreira* 1035–1076 < CAPRARIA (Rasico 2006, 102), *Lobeira* late 12th-early 13th cent. < LUPARIA (Russell-Gebbett 1965, 90–91); *dedebrei* ca. 1082–1131/*dezebrei* 1110 < *DECIPERE-HAIO (Rasico 2006, 103; Miquel Rosell 1945, vol. 1, 364), *engannarei* (id.) < *INGANNĀRE-HAIO; *Olla Freita* 1044 /*Ola Freita* 1132 < OLLA FRACTA (Rasico 1993, 68.), *forsfeit* 1048 < FORIS-FACTU (Russell-Gebbett 1965, 77), *feit* 12th cent. < FACTU (Rasico 1993, 115).

In the earliest Catalan documents the stressed open mid-vowels /*ɛ*/ (< E) and /*ɔ*/ (< O) appear continued as /*i*/ and /*u*/ respectively when preceding a yod, that is the palatal glide [j] which in the case of O > /*u*/ is occasionally preserved, as in LECTU > *llit* 'bed', SEX > *sis* 'six', MEDIU > *mig* 'half'; OC(U)LU > *ull* 'eye', PODIU > *puig* 'hill', NOCTE (> Old Cat. *nuit*) > *nit* 'night', etc. This development appears to have resulted from the diphthongization of the stressed open mid-vowel before yod in the preliterate period, similar to that which occurred in Occitan. However, in contrast to the latter, in

Catalan the diphthongs produced were reduced to the simple high vowels evidently through a process involving the assimilation of the open vowel to the surrounding glide elements; approximately [ɛj] > [ʲɛj] > [ʲej] > [ʲij] > [i] and [ɔj] > [wɔj] > [woj] > [wuj] > [u(j)] (Coromines ²1974, 248–250; Rasico 2006, 83–94).

The Act of Consecration of the cathedral at la Seu d'Urgell (ca. 860–880) offers two examples of /ɔ/ + [j] > /u/: *Muiopulto* < MODIU PULTU (Baraut 1978, 51), *Puioregis* < PODIU RĒGIS (Baraut 1978, 52). Additional examples of this development, together with others of /ɛ/ + [j] > /i/ are *Puio retundo* 1041 < PODIU ROTUNDU (Rasico 1993, 66), *pujos* 1048 < PODIOS (Russell-Gebbett 1965, 77), *ui* 1035–1076 < HODIE, *nuiti* id. < NOCTE (Russell-Gebbett 1965, 64); *lito* 1037 < LECTU (Rasico 1993, 66), *Castelvillo* 1035–1076 < CASTELLU VECLU (Rasico 2006, 90), *Valira* 1055 < VALERIA (Rasico 1993, 66).

3.1.2 Unstressed vowels

The loss of a word-final unstressed mid-vowel (/e/, /o/), except in cases where a final vowel was required for reasons of syllable structure, is another phenomenon that appears to have been completed very early in the preliterate period perhaps as a result of the influence of a Germanic superstrate (Visigothic or Frankish) upon locally spoken Latin (Rasico 2006, 143–144; Moran i Ocerinjauregui 1994, 19–20). Once again the act of consecration of the cathedral at la Seu d'Urgell (ca. 860–880) provides evidence of this development as the following toponyms reveal: *Kasamuniz* < CASA MENNĪCI or possibly MANNĪCI (Baraut 1978, 52), *Malangez* < *MALANGEICI (Baraut 1978, 52), *Linars* < LĪNĀRES (Baraut 1978, 51), *Kabrils* < CAPRĪLES, *Cortalz* < CORTE-ALIKIS, *Argilers* < ARGĪLLARIOS (Baraut 1978, 51). Written examples of the loss of word-final, unstressed -e and -o become particularly frequent as of the 10th century: *Cavallar* ca. 950 < CABALLĀRE, *linar* id. < LĪNĀRE, *Oriol* id. < AUREOLU (Mundó 1974, 239–240), *Tres Mals* ca. 1014 < TRĒS MALOS (Rasico 2006, 142), *frument* ca. 1060 < FRUMENTU (Mundó 1974, 241), *obac* 1069 < OPĀCU (Baraut 1978, 147), *Olers* 1076 < OLLARIOS (Russell-Gebbett 1965, 72), *Franculi* id. < FRANCOLĪNU (Russell-Gebbett 1965, 72–73), *cam* 1085 < CARNE, *porc* 11th cent. < PORCU, *ort* id. < HORTU (Rasico 2006, 284).

In a post-tonic word-internal syllable Latin A (/a/) is regularly continued in Catalan as e (/e/), which is realized phonetically as [e] in the modern northwestern and southern dialects and as [ə] in the eastern dialects (but [a] in the Catalan dialect of l'Alguer on the island of Sardinia (Rasico 1993, 15–19; 2006, 119–127)). This change appears to have occurred due to the weakening of post-tonic vowels in the spoken Latin of Catalonia or at the very latest early in the preliterate period. As a result of the historical, systematic raising of post-tonic internal /a/ to /e/ the language presents forms and morphological alternations of the types: *ànec* 'duck' < ANATE, *espàrrec* 'asparagus' < ASPARAGU, *rave* 'radish' (< anc. *raven*) < RAPHANU, *cases* 'houses' <

CASAS (cf. *casa* < CASA), *bones* ‘good’ (fem. pl.) < BONAS (cf. *bona* < BONA), *cantes* ‘you sing’ (fam.) < CANTAS (cf. *canta* < CANTA/CANTAT), *canten* ‘they sing’ < CANTANT, *cantàvem* ‘we sang’ < CANTĀBAMUS (cf. *cantava* < CANTĀBAM/CANTĀBAT), *temien* ‘they feared’ < TIMĒBANT, *digues* ‘say’, ‘tell’ (imp.) < DĪCAS, etc.

Evidence of the raising of post-tonic A (> /a/) > /e/ is discovered in Catalan documents from at least the 10th century; and it is likely that its lack of appearance at an earlier date (perhaps in the 9th century when it probably was already a feature of popular speech) was due primarily to the influence of written Latin together with conservative scribal traditions. Two possible 9th century examples of -AS > -es appear in the toponyms *Cabanes* 854 < CAPANNAS (Guiter 1960, 37) and *Kannes* 898 < CANNAS (Udina Martorell 1951, 120). Additional forms containing the same change are: *Cabages* 960 < CAPANNAS (Rasico 2006, 124), *Nogere* ca. 995 < *NOCARIA (Rasico 1993, 19), *volies* 1010–1035 < VOLĒBAS (Miret i Sans 1912, 382), *Ribeles* 1035–1079 < RĪPELLAS (Russell-Gebbett 1965, 65), *Vigoles* 1085 < VĪNEOLAS, *pernes* 11th cent. < PERNAS (Rasico 1993, 19), *podien* 1115 < *POTĒBANT (Miret i Sans 1908, 17), *Vilamigane* 1179 < VĪLLA MEDIĀNA (Rasico 1993, 74), *avien* 12th cent. < HABĒBANT (Pujol 1913, 15).

The raising of /a/ to /e/ (perhaps pronounced [ɛ]) in the third-person singular of verbs from the northwestern Catalan area is suggested by forms such as *proclame* 1036–1079 < PROCLĀMAT (Russell-Gebbett 1965, 65), *esbalczave* 1085 < *EXBALTIĀBAT (Rasico 1993, 19), *ere* 1080–1095 < ERAT (Russell-Gebbett 1965, 75).

Prior to the 13th century little written evidence is found that would indicate a confusion of atonic /a/ and /e/ (> [ə]) as occurs today in unstressed syllables in the eastern Catalan dialects. A number of forms appearing in medieval documents that would seem to suggest such a confusion prior to the 13th century are, for the most part, attributable to vocalic dissimilation, the influence of an adjacent consonant or simply to scribal error, as in *enap* 986 < HANAPPU, *exada* id. < ASCIĀTA (Russell-Gebbett 1965, 61), *enar* 12th cent. < AMNĀRE (Russell-Gebbett 1965, 83). A similar observation may also be made regarding unstressed /o/ and /u/ which, in the greater part of the eastern Catalan area, have become confused in speech as [u] in both pre-tonic as well as post-tonic position. Instances of /o/ > /u/ (i.e. *u* for expected *o*) found prior to the late 15th or early 16th century are due in general to the assimilation of /o/ to a following stressed high vowel (generally /i/), as in *ruuira* ca. 950 < *ROBERIA (Mundó 1974, 240), *cunil* 1040 < CUNICULU (Baraut 1978, 129).

3.2 Development of Consonants

Due to considerations of space, in this section only a selected number of consonantal developments will be examined for which there is evidence found in documents from the preliterate period. For a general analysis of the evolution of the Catalan system of consonants one may consult Badia i Margarit (1981a, 177–253), Moll (1991, 90–119),

Batlle et al. (2016, 85–169) and Recasens i Vives (2017, 169–404). Russell-Gebbett (1965, 30–40) studies the evolution of consonants in medieval Catalan based upon the texts edited in his volume, most all of which are primary sources for the study of Old Catalan.

3.2.1 Bilabials and labio-dentals

In a large area of the Catalan-speaking territory the reflexes of the intervocalic Vulgar Latin bilabial occlusive consonants *B-*, *-P-* developed to a voiced bilabial stop (/b/, spelled *b*), while intervocalic *-B-* and *-F-*, and *V* in any position initially gave the voiced bilabial fricative /β/ (spelled *v*), as in *BUCCA* > *boca* ‘mouth’, *RĪPA* > *riba* ‘riverbank’, *FABA* > *fava* ‘bean’, *RAPHANU* (> anc. *raven*) > *rave* ‘radish’, *VACCA* > *vaca* ‘cow’, *COVA* > *cova* ‘cave’, etc. At some point, and perhaps still in spoken Latin, /β/ modified its point of articulation in some areas from bilabial [β] to labio-dental [v], which subsequently became phonologized as /v/, a change that must have occurred well before (*B-*, *-P-* >) /b/ developed the fricative allophone [β]. The contrast between /b/ and /v/ is retained today in the Catalan of the Balearic Islands, l’Alguer, most of Valencia (with the exception of the *Apitxat* subdialectal region of central Valencia), and sporadically in some areas of southeastern Catalonia (the Camp de Tarragona, Baix Penedès, the Priorat and the Ribera d’Ebre). Elsewhere /b/ and /v/ have merged as /b/ with the allophones [b]/[β] (Badia i Margarit 1981a, 178–179, 184–186, 190, 192; Moll 1991, 90–91, 94; Gulsoy 1993, 128–131).

The confusion of /b/ and /v/ in Old Catalan does not appear to have occurred until the late 13th century at the very earliest; and it is probable that the merger of the two phonemes was not completed until sometime after the Reconquest and repopulation of the Balearic Islands (1229–1287) and Valencia (1238) and most likely also following the Catalan occupation of l’Alguer (1354). However, a number of examples are found in preliterary Catalan documents that suggest the confusion of *b* and *v*, a fact that may point to an incipient loss of contrast between the two voiced consonants in word-initial and intervocalic positions. Nevertheless, given the exceptionally early date of some of the aforementioned examples, most may simply be due to individual scribal errors or to factors such as assimilation, dissimilation, etc.

Among the possible early examples of the confusion of *b* and *v* are: *vona voluntate* 857 < *BONA VOLUNTĀTE* (Baraut 1978, 59), *Vergitano* 885 < **BERGITĀNU* (Rasico 2006, 159), *bindo* 892 < *VINDO*, *binea* id. < *VĪNEA* (Russell-Gebbett 1965, 57), *balle* 906 < *VALLE* (Rasico 2006, 159), *Castelbel* 1063 < *CASTELLU VECLU* (Miquel Rosell 1945, vol. 1, 238).

When the Old Catalan phoneme /β/ became syllable-final or word-final as the result of a vocalic syncope or apocope, it vocalized to *-u* ([-w]) perhaps already by the 9th century, as the following examples suggest: *Carcoude* 839 (*OnCat*, vol. 3, 266, s.v.

Carcolze) *Carcolde* 964 < *KARRI-KÓ-BIDE;² *auré* 1010–1018 < *HABÈRE-HAIO (Miret i Sans 1912, 381); *uius* 1041–1075 < VĪVOS (Miret i Sans 1912, 383), *deu* 11th cent. < DĒBET (Russell-Gebbett 1965, 71), *bou/bous* 12th cent. < BOVE(S), *aus* 12th cent. < AVES (Rasico 1993, 117).

3.2.2 Voiced dental fricatives and affricates

In early Old Catalan the reflexes of -D- (> /ð/) and -TJ-, -C^{e,i} (> /dʒ/) merged as a weak voiced dental fricative (/ð/) which frequently appears represented in documents as *d*. In pre-tonic position this phoneme was generally eliminated, as in CRŪDĒLE > *cruel* ‘cruel’, LAUDĀRE > *lloar* ‘to praise’, RATIŌNE > *raó* ‘reason’, PRETIĀRE > *prear* ‘to value’, VICĪNU > *vei* ‘neighbor’, PLACĒRE > *plaer* ‘pleasure’, etc. In post-tonic position, however, /ð/ > /z/ between vowels therefore merging with the voiced alveolar result of -S-, -NS- (> /z/) as in ALAUDA > *alosa* ‘lark’, *INCLŪDINE > *enchusa* ‘anvil’, CRETEA > *gresa* ‘clay’, MALITIA > *malesa* ‘evil’, etc. (cf. CAUSA > *cosa* ‘thing’, *PREHĒNSA > *presa* ‘seizure’, etc.). In syllable-final and word-final positions /ð/ vocalized to -u (-[w]) evidently through a gradual process of velarization and labialization (Coromines 1974, 253–259; Gulsoy 1993, 141–156; Rasico 2006, 164–197).

Catalan documents provide evidence suggesting that the deaffrication of /dʒ/ (< -TJ-, -C^{e,i}) and its merger with /ð/ (< -D-) as the weak dental fricative /ð/ must have taken place very early in the medieval period. Among numerous examples that would appear to support this change are: *Paladolo* 979 < PALATIOLU (Guiter 1960, 28.), *Planedes* id. < PLANITIAS (Guiter 1960, 26), *Avida* 983 < AVITIĀNU (Baraut 1978, 103), *Maradanus* 990 < MARETIĀNUS (Guiter 1960, 8), *Lidinia* 1035–1076 < LĪCINIA, *franche-das* 1062 < *FRANKITIAS (Rasico 2006, 182). Some examples of /ð/ in syllable-final and word-final position (the latter due to apocope) are *Paladdanum* 833 < PALATIU-DANI (Alart 1877, 116), *Dalmad* 1063 < DALMATIU (Rasico 2006, 182), *fed* 1085–1095 < FĒCIT (Miret i Sans 1908, 12).

Loss of /ð/ (< -TJ-, -C^{e,i}, -D-) in pre-tonic position is seen in *Marreans* 997 < MARETIĀNUS (Guiter 1960, 8), *Palaol* 1098 < PALATIOLU (Baraut 1978, 164), *Beluer* 12th cent. < BELLU-VIDĒRE (Pujol 1913, 14), *faie* id. < FAC(I)ĒBAT (Russell-Gebbett 1965, 82); and in a post-tonic syllable there appear *Planées* 1162 < PLANITIAS (Rasico 2006, 184), *folees* 12th cent. < *FOLLITIAS (Pujol 1913, 13), *francea* id. < *FRANKITIA (Pujol 1913, 15).

The vocalization of syllable-final and word-final /ð/ to -u (-[w]) is attested to by the following forms: *Palaldano* 1017 and *Palau dano* 1154 < PALATIU-DANI (Alart 1877,

² The same toponym appears in a document from 964 as *Carcolde* (Guiter 1960, 49) in which the implosive -u has been replaced by its near acoustic equivalent -l: *KARRI-KÓ-BIDE > *Carcobde > *Carcoude* > *Carcolde* (> mod. *Carcolze*).

116),³ *leuda* 1061 < LICITA (Miquel Rosell 1945, vol. 2, 101), *aloudem* 943, *aloudio* 1064, *aloudium* 1072 < ALŌD (Alart 1877, 124),⁴ *ociure* 12th cent. < OCCĪDERE (Pujol 1913, 14).

3.2.3 Deaffrication of Old Catalan /ts/ and merger with /s/

The voiceless dental affricate /ts/ of Old Catalan, which derived from C^{e,i}- in word-initial and post-consonantal position, as well as from intervocalic -CJ- and post-consonantal TJ-, contrasted with the voiceless alveolar fricative /s/ (from word-initial S-, post-consonantal and word-final -s, and intervocalic -SS- and -RS-) until well into the preliterate period when they merged as alveolar /s/. Although there is some evidence found in written sources that appears to indicate the beginning of a loss of contrast between these two sibilants from as early as perhaps the late 10th century, via the deaffrication of /ts/ > /s/, a merger of the two phonemes was not likely to have been completed until approximately the 13th century, a change which seems to have spread gradually in a northwesterly direction from the Mediterranean coast (Coromines 1974, 297; 1976–1977, vol. 1, 14–29). Some early forms that may indicate an initial confusion of Old Catalan /ts/ with /s/ are: *Balbeses* 914 < VALLE-^{*}BIKIAS? (Rasico 2006, 187),⁵ *pessas* 992 (Rasico 2006, 188), *pesas* 1030 (Coromines 1976–1977, vol. 1, 25) < ^{*}PETTIAS, *Massana* 1009 < ^{*}MATTIĀNA (Rasico 2006, 188), *Flassa* 1041 < FLACCIĀNU (Guiter 1960, 30).

3.2.4 Assimilation and simplification of intervocalic consonant groups /mb/, /nd/, /ld/

In the very earliest of Catalan texts one discovers the simplification, through a process of progressive assimilation, of the groups -MB- (> /mm/) > /m/, -ND- (> /nn/) > /n/, and with less frequency and consistency -LD- (> /ll/) > /ul/, the latter solution resulting from the vocalization of syllable-final /l/. Examples of these changes are COLUMBU > *colom* ‘dove’, CAMBA > *cama* ‘leg’, MUNDU > *món* ‘world’, MANDĀRE > *manar* ‘to order’, BISULDŪNU (> ^{*}Besollún > ^{*}Besoulú > ^{*}Besolú) > *Besalú*, CALDARIA > anc. *callera* ~ *caulera* ‘boiling-pot’ (Badia i Margarit 1981a, 204–205; Moll 1991, 105; Coromines 1974, 214–216).

³ Guiter (1960, 49) cites *Palaudan* (967); however it is unclear if this form appears in an original document or is retranscribed in a text from a later period.

⁴ Coromines (*DECat*, vol. 1, 225, s.v. *alou*) explains forms such as *alaude*, *aloude*, *aloudio*, etc. as a written compromise between Late Latin ALODE and Catalan *alou* in which final /ð/ has become -u (-[w]).

⁵ Coromines (*OnCat*, vol. 2, 491, s.v. *Bessó*) mentions the village of *Torrebeses* (les Garrigues) that appears documented as *Torres Beces* (1359).

The act of consecration of the cathedral at la Seu d'Urgell (ca. 860–880) offers two examples of the assimilation and reduction of -MB- to /m/: *Palomera* < PALUMBARIA and the corresponding diminutive *Palomerola* < PALUMBARIOLA (Baraut 1978, 52). However, at an even earlier date there appears *coma* 815 < CUMBA (Guiter 1960, 38), a form that is also found in numerous documents from the 10th and 11th centuries: *comas* 927 < CUMBAS (Udina Martorell 1951, 227), *coma* 966 (Rasico 2006, 203), *comma* 976, 1088 (Udina Martorell 1951, 356; Guiter 1960, 38), *Chomalada* 11th cent. < CUMBA-LATA (Rasico 2006, 307); *Paluma* 913 < PALUMBA (Udina Martorell 1951, 159), *Colomera* 966 < COLUMBARIA (Rasico 2006, 203).

Examples of -ND- > /(n)n/ discovered in early Catalan texts include: *sponna* 899 < SPONDA (Baraut 1978, 67), *apennicio* 908 < APPENDICIO (Bastardas 1995, 117), *conamina* 967 and *conamina* 1094–1110 < CONDAMĪNA (Miret i Sans 1912, 350; Russell-Gebbett 1965, 69), *Annorra* 1003 < ANDORRA (Miret i Sans 1912, 351), *lanna* 1051 < LANDA (Miret i Sans 1912, 356), *atenniems* 1060 < ATTENDĒBAMUS (Russell-Gebbett 1965, 77).

Although examples of -LD- (> /ll/) > /ul/ are less frequent than the simplification of -MB- (> /mm/) > /m/ and -ND- (> /nn/) > /n/, the following are illustrative of the transformations undergone in popular speech by the Latin group -LD- and by the same group when it resulted from an early syncope (-L(I)D- > -LD-): *Bisillu* 1010–1018 < BISULDŪNU (Miret i Sans 1912, 381), *Bisullunense* 866, 1035–1076 < BISULDŪNENSE (Guiter 36; Rasico 2006, 204), *Arnall* ~ *Arnal* 1080–1095 < ARNALDU (Russell-Gebbett 74), *callera* 1157 < CALDARIA (Miret i Sans 1912, 393).

3.2.5 Evolution of two distinct Old Catalan palatal laterals

In Northwestern Catalan, the Valencian region, and some areas of the eastern Catalan dialectal territory the reflexes of the groups -LJ-, -C'L-, -G'L-, together with those of -LL- and word-initial L-, have become confused as the voiced palatal lateral /ʎ/ (spelled *ll*) in the standard literary language, as in *FĪLIU* > *fill* 'son', *FOLIA* > *fulla* 'leaf', *PALEA* > *palla* 'straw', *APIC(U)LA* (> *APIC'LA*) > *abella* 'bee', *VET(U)LU* (> *VEC'LU*) > *vell* 'old', *RĒG(U)LA* (> *RĒG'LA*) > *rella* 'ploughshare', *STRAG(U)LU* (> *STRAG'LU*) > *estrall* 'damage', etc.; *CABALLU* > *cavall* 'horse', *AMPULLA* > *ampolla* 'bottle', *VALLE* > *vall* 'valley', *LŪNA* > *lluna* 'moon', *LACTE* > *llet* 'milk', *LOCU* > *lloc* 'place', *LINGUA* > *llengua* 'tongue', 'language', etc.

In an extensive part of eastern Catalonia located approximately between the Ter and the Llobregat rivers (excluding the city of Barcelona and its immediate vicinity), as well as in the Balearic Islands, a contrast exists between the modern reflexes of what must have been two distinct types of palatal lateral phonemes in Old Catalan: one originating from the groups with a glide (-LJ-, -C'L-, -G'L-) and the other from an intervocalic double lateral (-LL-) or a word-initial lateral (L-). In those areas where the two types of old palatal laterals continue to be distinguished the contrast is between

/j/ (< -LJ-, -C'L-, -G'L-) and /ʎ/ (< -LL-, L-) as the following examples reveal: *agulla* [ə'ɣujə] 'needle' (< ACŪC(U)LA), *ull* ['uj] 'eye' (< OC(U)LU), *tell* ['tej] 'linden' (< TILIU); *castell* [kəs'teʎ] 'castle' (< CASTELLU), *sella* ['seʎə] 'saddle' (< SELLA), *poll* ['poʎ] 'chick' (< PULLU); *llac* ['ʎak] 'lake' (< LACU), *llebre* ['ʎeβrə] 'hare' < LEP(O)RE, etc. (Badia i Margarit 1981a, 49–66; Veny 1991, 20–28).⁶

Old Catalan appears to have possessed two distinct palatal lateral phonemes whose contrast may have been based upon a difference in their degree of palatalization: the common reflex of -LJ-, -C'L-, -G'L-, which frequently appears represented in medieval texts by *yl* and *yll* among other solutions (*lg*, *il*, *ill*, etc.), may have been more fully palatal in quality (i.e., perhaps with respect to its point of articulation – palatal or prepalatal: [jʎ] or [jʎj]) than the result of -LL- and L-. The realization of the latter phoneme, which is frequently represented as *l* or *ll* between vowels and nearly exclusively as *l* at the beginning of a word until nearly the end the Middle Ages, may initially have been more alveolar (retroflex?) than that of /jʎ(j)/ and perhaps only later became more fully palatalized once /jʎ(j)/ had begun to shift its manner of articulation in some areas from a palatal lateral to a palatal fricative (/j/). According to Coromines (1976–1977, vol. 1, 41): 'The only thing certain is that in the Middle Ages *yl* must have had a stronger palatal quality than that of *ll*, since that is what the modern pronunciation of *y* clearly confirms' (translation PDR).

The unconditioned delateralization and fricativization of the old phoneme /jʎ(j)/ (> /j/) is discovered in Catalan documents from at least the 11th century: *Corneia* 1054 < CORNELIĀNU (Miquel Rosell 1945, vol. 1, 267), *pareiada* ca. 1065 < *PARIC(U)LĀTA (Rasico 2006, 220), *reia* 1066 < RĒG(U)LA (Russell-Gebbett 1965, 70), *Toreió* ca. 1075 < TAURELIŌNO, *rostoi* id. < *RUSTUC(U)LU (Rasico 2006, 307), *oueia* ca. 1094 < OVIC(U)-LA (Rasico 2006, 220), *Casteion* 1128 < *CASTELLIŌNE (Guiter 1960, 33).

Evidence of the confusion of (-LJ-, -C'L-, -G'L- >) /jʎ(j)/ with (-LL- >) /ʎ/ appears at approximately the same period, especially in texts from northwestern Old Catalonia:⁷ *Pallerols* 967 < *PALEARIOLOS (Miret i Sans 1912, 350), *uermellas* 1058 < VERMIC-(U)LAS (Miret i Sans 1912, 384), *talas* 11th cent. < TALEAS, *chonilos* id. < CUNĪC(U)LOS, *Salent* id. < SALIENTE (Rasico 2015, 151), *parellades* 1152 < *PARIC(U)LĀTAS (Miret i Sans 1908, 14), *collons* 12th cent. < COLEŌNES (Pujol 1913, 14).

⁶ The pronunciations indicated are those of the Eastern Catalan dialect. On the history of the two types of palatal lateral consonants in Old Catalan, see Coromines (²1974, 289–292) and especially Coromines (1976–1977, vol. 1, 13–85). Rasico (2006, 218–224) provides a detailed analysis, with supporting documentation, of the evolution of these two lateral phonemes in Old Catalan.

⁷ According to Coromines (1976–1977, vol. 1, 18): 'My study of the old texts has shown that this distinction [jʎ(j)]:/ʎ/ was observed, in the 12th and 13th centuries, in a much larger area which covered the greater part of the Principality, including, much more than the other, the well-populated regions, except for the Alt Empordà, Pallars, the Ebro Valley excluding the Segre, and a narrow coastal strip which from the mouth of the Ebro extended right up to Barcelona. However in subsequent centuries this isogloss gradually retreated toward the east, until it came to coincide with the present limit' (translation PDR). See Coromines' Mapa 1 (ibid.), opposite p. 108.

As noted previously, there is scarce written evidence of the palatalization of word-initial *L-* until the late Middle Ages and beyond this period.⁸ Nonetheless, an adventitious example of *L- > ll-* occurs in the toponym *Sallagosa* ca. 860–880 ‘area of willows’ < SALICŌSA, in which the initial syllable *Sa-* appears to have been falsely analyzed in popular speech as the article derived from IPSA (> *sa*), as if it were IPSA *LICŌSA (> *Sa Llegosa > Sallagosa*) (Baraut 1978, 51; Coromines 1976–1977, vol. 1, 73; *OnCat*, vol. 7, 20, s.v. *Sallagosa*).

Additional early examples that suggest the palatalization of the word-initial lateral are *Llicinia* 1001 < LICĪNIA and *Lloroi* 1004 < *LORRUI (Coromines 1976–1977, vol. 1, 72).

3.2.6 Loss of a word-final alveolar nasal

An especially distinctive phonological development of Old Catalan was the loss of the word-final alveolar nasal /n/ whether of Latin or Romance origin. Although this phoneme was retained in some forms (frequently due to syntactic-phonetic factors, as in *bon home* ‘good man’ < BONU HOMINE, *ben fet* ‘well done’ < BENE FACTU, *algun altre* ‘some other (one)’ < *ALIU-ŪNU ALTERU, etc.), it generally was eliminated in word-final position in words derived from Latin singular paroxytone forms while being retained in the plural (PANE > *pan* > *pa* ‘bread’, pl. PANES > *panes*; PLANU > *plan* > *pla* ‘flat’, pl. PLANOS > *plans*; GERMĀNU > *germàn* > *germà* ‘brother’, pl. GERMĀNOS > *germans*; MANU > *man* > *mà* ‘hand’, pl. MANUS > *mans*, etc.).⁹ This loss of word-final /n/ was extended as well to words derived from Latin singular proparoxytones and, eventually, also to their plurals in some dialects, as in HOMINE > *homen* > *home* ‘man’, pl. HOMINES > *hòmens* > *homes*; TERMINU > *termen* > *terme* ‘term’, pl. TERMINOS > *tèrmens* > *termes*; etc. Northwestern Catalan and Valencian, however, generally retain the /n/ in the plural of paroxytones derived from Latin proparoxytones: HOMINES > *hòmens*, TERMINOS > *tèrmens*, MARGINE > *màrgens*, etc. The word-final /n/ was also retained by analogy in some third-person singular verb forms as well as in a number of learned words: PONIT > *pon* ‘lays’; REMĀNET > *roman* ‘remains’; EXĀMEN > *examen*

⁸ Coromines (1976–1977, vol. 1, 62–63) observes that ‘As late as the 17th century there is still a text in which the spelling *l-* for *ll-* is almost general, and some others in which it predominates [...]. In the majority of those from the 17th century *ll-* is nearly general, and in the 16th century very predominant, although in certain manuscripts from the first third of the 16th century *l-* still systematically appears [...]. It hardly needs mentioning, therefore, that in all of the medieval manuscripts *l-* is general or at least very predominant, at least in literary manuscripts. But there can be no doubt that, already in the Middle Ages, while writing *l-*, one pronounced *ll-*’ (translation PDR).

⁹ The Roussillonese dialect, however, generally eliminates the final /n/ before the plural morpheme /s/: PANE(S) > *pa/pas*, RATIOŅE(S) > *raó/raós* (pronounced [rə’u]/[rə’us], etc. (Coromines ²1974, 303–304, 313–314).

‘examination’; CERTĀMEN > *certamen* ‘contest’, etc (Badia i Margarit 1981a, 241–242; Rasico 2006, 238).

The loss of Romance-final /n/ is seen in Catalan texts from approximately the mid-10th century, a fact that suggests that it already may have begun to be eliminated in popular speech by the late-9th century. Some early examples of the loss of -/n/ are *Uuigmara* ca. 950, *Wimara* mid-10th cent. < WIGMARĀNE (Mundó 1974, 239; Rasico 2006, 245); *Auida* 983 < AVITIĀNU (Guiter 1960, 18), *té* 1036–1079 < TENET (Russell-Gebbett 1965, 65), *ui* 1037 < VĪNU (Miret i Sans 1912, 353), *Franculi* 1076 < FRANCOLĪNU (Russell-Gebbett 1965, 72), *pla* 1095 < PLANU (Miret i Sans 1912, 388), *Lordá* 12th cent. < LURRI-TE-ĀNU, *Montfalcó* 12th cent. < MONTE-FALCŌNE (Rasico 2015, 152).

While most early medieval Catalan texts retain the /n/ before the plural morpheme /s/ as in modern Northwestern Catalan dialects and in Valencian, there are, nonetheless, some examples of the loss of the nasal during this period among which are *ses* 1035–1076 < *SINES (Russell-Gebbett 1965, 64) and *ioues* 12th cent. < IUVENES (Pujol 1913, 13).

4 Morpho-syntactic developments

4.1 General observations

Little can be said about the morpho-syntactic features of early Old Catalan until approximately the 11th and 12th centuries, since prior to this period documents were written in Late Latin with only the occasional appearance of a Romance form: generally toponyms, personal names or other terms often thinly disguised as Latin. It is only as a consequence of profound changes in Catalan society (i.e. the institutionalization of the feudal system) that occasional phrases and even entire texts appear that are written fully or for the most part in Catalan (Moran i Ocerinjauregui 1994, 89–90; Moran/Rabella 2015, 119–122; Rasico 2015). This circumstance, together with the nature of the documents themselves (feudal oaths and agreements; documents expressing grievances for the abuse of feudal rights, etc.), are among the principal factors responsible for a limited and often repetitious lexicon, as well as for relatively simple syntactic structures, that characterize the texts described.

4.2 Nouns and adjectives

4.2.1 Nominatives

Few vestiges of the Latin declensional system are found in Old Catalan documents. An exception to this, however, is the pronoun system which will be examined below (§ 4.3).

Some examples of the retention of the nominative case in both nouns and adjectives are: *Petrus Arnals* 11th cent. < PETRUS ARNALDUS, *inimics* 11th cent. < INIMĪCUS (Rasico 2015, 153), *parelos* 1085 < *PARICULUS (Rasico 2006, 288), *fideles* 1043–1098, *fidels* late 11th cent., *fedels vos seré* 1131 < FIDĒLIS (Russell-Gebbett 1965, 40, 71, 79), *om* late 11th cent. < HOMO (Russell-Gebbett 1965, 71).

4.2.2 Definite articles and demonstratives

The definite articles of Old Catalan derive from two distinct Latin demonstratives: ILLE and IPSE. The former, in its accusative form ILLU, gives the articles *lo* (reduced to *l* following a word ending in a vowel, and especially after the prepositions *a* < AD and *de* < DE, as well as before a word beginning with vowel), pl. *los* < ILLOS; while the latter, also in its accusative form IPSU, gives *so* (reduced to *s* as in the preceding case), pl. *sos* < IPSOS. The corresponding feminine articles are *la* < ILLA (reduced to *l* before a word beginning with a vowel), pl. *les* < ILLAS; and *sa* < IPSA (reduced to *s* before a vowel), pl. *ses* < IPSAS. Some examples of these articles including their reduced forms are: ILLU PANE > *lo pa* ‘the bread’, PORTAT ILLU PANE (> *porta lo pa*) > *porta’l pa* ‘he takes the bread’, AD ILLU PANE (> *a lo pa*) > *al pa* ‘to/for the bread’, DE ILLU PANE (> *de lo pa*) > *del pa* ‘of the bread’, ILLOS PANES > *los pans* (> *els pans*); ILLA TABULA > *la taula* ‘the table’, ILLA AMĪCA (> *la amiga*) > *l’amiga* ‘the friend’ (f.), ILLAS AMĪCAS > *les amigues* ‘the friends’ (f.); IPSU PANE > *so pa* ‘the bread’, PORTAT IPSU PANE (> *porta so pa*) > *porta’s pa* ‘he takes the bread’, AD IPSU PANE (> *a so pa*) > *as pa* ‘to/for the bread’, DE IPSU PANE (> *de so pa*) > *des pa* ‘of the bread’, IPSOS PANES > *sos pans* (> *ets pans*); IPSA TABULA > *sa taula* ‘the table’, IPSA AMĪCA (> *sa amiga*) > *s’amiga* ‘the friend’ (f.), IPSAS AMĪCAS > *ses amigues* ‘the friends’ (f.), etc.

In Modern Catalan the articles derived from the Latin accusative of ILLE have become general in the standard language, while those from IPSE survive in popular speech in the Balearic Islands as well as in a small area along the Catalan coast from approximately the town of Cadaqués on the Cap de Creus (Alt Empordà) in the north to that of Blanes (la Selva) in the south. The articles derived from IPSE are also found fossilized in numerous Catalan toponyms (e.g. *Espui* < IPSU PODIU, *Sant Joan Despí* < DE IPSU PĪNU, *Sant Martí Sarroca* < IPSA *ROCCA, *Sant Climent Sescebes* < IPSAS CIPAS, etc.) even in areas where today the articles from ILLE are used. In addition, in Northwestern Catalan the medieval forms of the latter are preserved in the

masculine articles (*lo ~ l, los*) (Badia i Margarit 1981a, 314–321; Moll 1991, 139–140; Veny 1991, 48–49, 65, 84–85, 134).

Some examples of the definite articles in Catalan documents from the period being considered are: *lo castel* 1085 < ILLU CASTELLU, *lo vi* 12th cent. < ILLU VĪNU, *e-l comendador* 12th cent. < ILLU COMMENDATŌRE, *del castel* 1085 < DE ILLU CASTELLU (Rasico 2006, 290), *lo termen* 1094–1110 < ILLU TERMINU (Russell-Gebbett 1965, 68), *la francea* 12th cent. < ILLA *FRANKITIA, *la tore* id. < ILLA TURRE (Rasico 2015, 155), *l'eglesia/l'esglesia* id. < ILLA *ECLĒSIA (Rasico 2015, 156), *los homens* id. < ILLOS HOMINES, *e-ls explets* id. < ET ILLOS EXPLICITOS (Rasico 2006, 290), *las casas* 12th cent. < ILLAS CASAS (Rasico 2006, 327).

Evidence of the article derived from IPSE is found in numerous preliterary Catalan documents where, in many cases, it functions as the determiner of a toponym: *Gerat dez Penal* 12th cent. < DE IPSU PINNĀLE, *na Maria de za Porta* id. < DE IPSA PORTA (Rasico 2015, 155), *P. de ça Font* late 12th cent. < DE IPSA FONTE (Rasico 2006, 323), *za Corit* 12th cent. < IPSA *LAK-GORRITZE (Russell-Gebbett 1965, 83). In other instances, however, there are examples that reveal the use of IPSE as an independent article, as in *zos porcs e-z bous* 12th cent. < IPSOS PORCOS ET IPSOS BOVES, *zo mas* id. < IPSU MANSU, *e pres z'aver seu* id. < IPSU HABĒRE, *pres-lo pez collons* id. < IPSOS COLEŌNES (Russell-Gebbett 1965, 83).

As in the case of the definite articles, there were competing demonstrative adjectives in early Medieval Catalan: those derived from ISTE, IPSE and ILLE in their respective accusative forms, the first two giving directly *est* (f. *esta*) ‘this’, *eix* (f. *eixa*) ‘that’ in the singular and, in the plural, generally reinforced by ECCE- or *ECCU- (*ECCU-ISTU >), *aquest* (cf. f. *aquesta*; pl. *aquests, aquestes*); *ECCU-IPSU > *aqueix* (cf. f. *aqueixa*; pl. *aqueixs, aqueixes*) ‘that’ (nearby); as was also in general the case for ILLE (> *ECCU-ILLU) > *aquell* (cf. f. *aquella*; pl. *aquells, aquelles*) ‘that’ (remote) (Badia i Margarit 1981a, 309–314; Moll 1991, 138).

Examples of the demonstrative adjectives discovered in early medieval Catalan documents include the following: *de esta hora* 1035–1076 < DE ISTA HŌRA (Russell-Gebbett 1965, 64), *ache(ce)st castel* 1085 *ECCU-ISTU CASTELLU (Rasico 2015, 290), *ad achel...o [a]d achella* 1043–1098 < AD *ECCU-ILLU/ILLA (Russell-Gebbett 1965, 66), *achelas alods* late 11th cent. < *ECCU-ILLAS ALODS (Russell-Gebbett 1965, 71). The demonstratives such as those in *celes veds* late 11th cent. < *ECCE-ILLAS VICES (Russell-Gebbett 1965, 71), *zel ome* 12th cent. < *ECCE-ILLU HOMINE, and *cels logs* late 12th-early 13th cent. < *ECCE-ILLOS LOCOS (Russell-Gebbett 1965, 83) are most likely due to Occitan influence (Russell-Gebbett 1965, 45).

4.2.3 Possessives

Old Catalan had two types of possessive adjectives corresponding to a single possessor: a set of full or tonic forms, normally preceded by a definite article, which could be

pronominalized; and a set of reduced atonic forms used without a definite article (Badia i Margarit 1981a, 305–309; Moll 1991, 137–138).

The former group, which derived from MEU(M), *TEU(M) (analogical to MEU; cf. Cl. Lat. TUUM), *SEU(M) (analogical to MEU; cf. Cl. Lat. SUUM), gave the masculine forms *meu* (MEOS > pl. *meus*) ‘my’, *teu* (*TEOS > pl. *teus*) ‘your’, *seu* (*SEOS > pl. *seus*) ‘his, her’; while the corresponding feminine possessives evolved regularly from MEA(M) > *mia* (MEAS > pl. *mies*) ‘my’, TUA(M) > *tua* (TUAS > pl. *tues*) ‘your’, SUA(M) > *sua* (SUAS > pl. *sues*) ‘his, her’, tonic forms that were replaced at an early date by the analogical *meua/meva* (pl. *meues/meves*), *teua/teva* (pl. *teues/teves*) and *seua/seva* (pl. *seues/seves*) respectively. The tonic adjectives corresponding to a plural possessor developed from NOSTRU(M)/NOSTRA(M) > masc. *nostre*/fem. *nostra* (NOSTROS/NOSTRAS > pl. *nostres*) ‘our’, VOSTRU(M)/VOSTRA(M) > masc. *vostre*/fem. *vostra*, (VOSTROS/VOSTRAS > pl. *vostres*) ‘your’; and in the third person *ILLÛRU(M) (for ILLÛRUM) produced *llur* (pl. *llurs*) ‘their’.

The atonic short possessives used with a single possessor derived from the popular unstressed forms *MUM > *mon* (pl. *mos*) ‘my’, *TUM > *ton* (pl. *tos*) ‘your’, *SUM > *son* (pl. *sos*) ‘his, her’ in the masculine, and, in the feminine, *MA > *ma* (pl. *mes*) ‘my’, *TA > *ta* (pl. *tes*) ‘your’, *SA > *sa* (pl. *ses*) ‘his, her’. For more than one possessor the forms used were those described in the preceding paragraph: masc. *nostre* (pl. *nostres*), fem. *nostra* (pl. *nostres*) ‘our’; masc. *vostre* (pl. *vostres*), fem. *vostra* (pl. *vostres*) ‘your’; and the common third-person *llur* (pl. *llurs*) ‘their’.

Examples of the stressed possessive adjectives of Old Catalan are: *ab meu consel* late 11th cent., *ab lo meu consellg* 1131 < MEU(M) (Russell-Gebbett 1965, 71, 79); *de la mia onor* 1080–1095 < MEA(M) (Russell-Gebbett 1965, 75); *ni de tua onore* 1035–1076 < TUA(M) (Russell-Gebbett 1965, 64); *et tuas cavalgadas* ca. 1088 < TUAS (Russell-Gebbett 1965, 78); *et ad sua mulier* 1085–1095 < SUA(M) (Russell-Gebbett 1965, 75); *per les sues oradures e per les sues folees* 12th cent. < SUAS (Russell-Gebbett 1965, 82); *nos fan nostros amigs entendre* 1190–1210 < NOSTROS; *la valença de vostra onor* 1072–1079 < VOSTRA (Russell-Gebbett 1965, 85); *per vostres missadges* 1072–1079, *de vostros quars* 1131 < VOSTROS (Russell-Gebbett 1965, 78–79); *tro lor dret* 1080–1095 < ILLÛRU(M) (Russell-Gebbett 1965, 75),¹⁰ *et sine e[n]gan lur* 1039–1050 < *ILLÛRU(M) (Russell-Gebbett 1965, 76).

Two examples of pronominalized possessives, both occurring in the same text, are: *e ag-ne tot lo seu* < ILLU *SEU(M), *per fer dez seu* 12th cent. < IPSU(M) *SEU(M) (Russell-Gebbett 1965, 83).

Unstressed possessive adjectives are seen in the following examples: *et a ma muler* 1080–1095, *quar desmentist ma mulier ante me* (id.) < *MA (Russell-Gebbett 1965, 74–75); *che no siria mos dons* 1085–1095, *mos homes ni mos valedors* 1211 < *MOS (Russell-Gebbett 1965, 74, 92); *senes ton engano* 1035–1076, *qui a ton corps se tenet* late

¹⁰ The use of *lor* suggests Occitan influence (see Grandgent 1905, 104).

11th cent. < *TUM (Russell-Gebbett 1965, 64, 71); *no-t tolrei ta vita ne ta membra* id. < *TA, *per tos missaticos* id. < *TOS (Russell-Gebbett 1965, 71); *che-t faça tes osts* ca. 1088 < *TAS (Russell-Gebbett 1965, 78); *con omo debet eser a son sengore* 1035–1076, *que en B. ne pert son dret* (12th cent.) < *SUM (Russell-Gebbett 1965, 64; Rasico 2006, 290); *e-l comanná a ssa muler* 1080–1095, *no-l jachí a ssa mort* 1190–1210 < *SA (Russell-Gebbett 1965, 74, 86); *a Felgera ab sos termens* 1085 < *SOS (Rasico 2006, 290).

4.3 Personal pronouns

As noted previously (§ 4.2.1), the Latin declension has been preserved partially in the Catalan personal pronoun system in which stressed forms generally derive from the Latin nominative case; while unstressed forms continue other cases, such as the accusative and dative, and function either as direct or indirect complements of a verb. Two additional pronouns with adverbial functions will be considered along with the latter group of forms (Badia i Margarit 1981a, 291–309; Moll 1991, 136–137; Russell-Gebbett 1965, 41–43).

4.3.1 Stressed pronouns

Subject pronouns derived from the Latin nominative are infrequent in early Old Catalan texts. However, their appearance becomes more common as of the 11th century in feudal oaths and documents expressing grievances (i.e. the *querimònies*, *clams* and *rancures* mentioned above in § 2). Some of the earliest examples of Catalan subject pronouns are *go/jo* < EGO: *go fileles vos ende sere* 1028–1047, *che jo no-l li doné* 1080–1095 (Russell-Gebbett 1965, 63, 74); *ço que jo a fer-le-n avia* late 11th cent. (Rasico 2015, 153); *tu* < TŪ: *che tu mi comonirás* 1035–1076, *a qui tu o dubtarás* 1043–1098 (Russell-Gebbett 1965, 64, 66); *vos* < VOS: *che vos mi comendaretes* 1043–1098, *qe vos me-n comonreds* 1072–1079 (Russell-Gebbett 1965, 64, 78); *nos* < NOS: *et si nos Guiellm et Azemar* 1060, *on nos som fort clamans* 12th cent. (Russell-Gebbett 1965, 77, 82); *el* < ILLE: *et el cholí-l* late 11th cent., *qu[e] el ne pagarie* id. (Rasico 2015, 153).

Pronouns functioning as prepositional complements in Old Catalan include: *mi* < MIHĪ: *si-s partia de mi* 1080–1095, *m'avia fait et dict a mi et a ma muler* id. (Russell-Gebbett 1965, 74); *ti ~ te* < TIBĪ: *a te Ragimundo* [...] *et a te Ermesende* 1028–1047, *et de te no-ns partirems* 1060, *per ti o per vestro misaticho* 1035–1076, *et si de ti dimens venerit* 1043–1098, *a ti Bertran* late 11th cent., *et alberg ab ti* (Russell-Gebbett 1965, 63–64, 66, 71, 77, 79); *el* < ILLE: *che ad el lo commanné* 1080–1095, *che no-l me fedés sens el* (Russell-Gebbett 1965, 74); *si* < SI: *ke no li forsfaça ad si neque ad sua honore* 1117 (Russell-Gebbett 1965, 69); *vos* < VOS: *a vos ende atenderé* 1028–1047 (Russell-Gebbett 1965, 63), *si de vos Ramonde chomte* 1035–1076 (Russell-Gebbett 1965, 64); *eles* < ILLAS: *entr[e] eles e-ls esplets* 12th cent. (Rasico 2015, 153).

4.3.2 Unstressed pronouns

Atonic (clitic) pronouns exhibit considerable polymorphism in Old Catalan documents as a consequence of their immediate phonetic environment, and they may appear in both proclitic and enclitic positions. Moreover, and unlike the modern language, in Old Catalan a direct object pronoun normally precedes an indirect object pronoun (or pronominal adverb) before a verb (Moll 1991, 136–137; Coromines ²1974, 274–275; Russell-Gebbett 1965, 42–43).

Representative examples of Old Catalan unstressed pronominal forms appearing in pre-verbal position include: *De ipssos castellos [...] go fideles vos ende seré* 1028–1047 < VOS INDE (Russell-Gebbett 1965, 63); *agutor la ti seré* 1035–1076 < ILLA TE (Russell-Gebbett 1965, 64); *quod Remon Gonball li-o jakescha e li-o defenescha per escrito* 1036–1079 < ILLI HOC (Russell-Gebbett 1965, 65); *de qual nudriment i fa hom* 1085 < IBĪ (Rasico 2006, 289); *que lo-i atenda* mid-11th cent. < ILLU IBĪ (Russell-Gebbett 1965, 68); *e no la vol redre* 1089–1095 < ILLA (Russell-Gebbett 1965, 75); *no me-n jachia re id.* < ME INDE (Russell-Gebbett 1965, 74).

Among examples of unstressed pronominal enclisis in early documents the following may be noted: *et reddet-li la conamina* 1094–1110 < ILLI; *et Remon comes receb-le-n ad homen* mid-11th cent., *et comanná-le-n Mir Arnal* 1080–1095 < ILLI INDE (Russell-Gebbett 1965, 68–69, 74); *e tengren-lo .III. dies pres* 12th cent. < ILLU; *e mes-les en pret id.* < ILLAS; *e menaren-lo a Cunquebela id.* < ILLU; *et preseren-s'i del feu de la vila id.* < SE IBĪ; *e portaren-lo-sse'n id.* < ILLU SE INDE (Rasico 2006, 289).¹¹

4.4 Verbs

As noted above, given the nature of the majority of documents under consideration in this section (primarily feudal oaths along with declarations of grievances and allegations of malfeasance), the personal forms of the verbs that appear are limited and not especially varied. Moreover, in view of the complexities and especially of the polymorphous nature of Old Catalan verb paradigms, due primarily to analogical developments, it is possible here to comment only upon a representative selection of verb forms found in the early documents.¹²

¹¹ On relative and interrogative pronouns in Old Catalan see Badia i Margarit (1981a, 321–324), Moll (1991, 140), Russell-Gebbett (1965, 45–47), and Coromines (²1974, 331).

¹² For a more complete overview of verbs in Old Catalan see Badia i Margarit (1981a, 329–376), Moll (1991, 144–179), and Coromines (²1974, 316–330).

4.4.1 Infinitive

The Vulgar Latin of Catalonia reduced the four Classical Latin verb conjugations (-ĀRE, -ĒRE, -ERE, -ĪRE) to three (-ĀRE > -ar, -ERE > -re, -ĪRE > -ir) generally by remodeling those of the second conjugation (-ĒRE) as third conjugation verbs (-ĒRE > -ERE) with only a small number of exceptions (e.g. DEBĒRE > *deure* ‘to owe’, VIDĒRE > *veure* ~ OCat. *veer* ‘to see’, TENĒRE > *tenir* ‘to have’, etc.).

Among examples of the infinitive in Old Catalan documents are: *que solie dar a Sen Pere* 12th cent. < DĀRE, *les conamines que ssolien laurar* id. < LABORĀRE (Russell-Gebbett 1965, 82), *entrar et ixir et geriar* 1055–1098 < INTRĀRE/EXĪRE/*guerra* (< WERRA) + -iar (< -IDIĀRE) (Russell-Gebbett 1965, 68), *et dreçar no-m volg* 1080–1095 < DIRECTIĀRE, *per gitar ni per metre* id. < *TECTĀRE/MITTERE (Russell-Gebbett 1965, 74), *ni deceber non ti faré* 1047–1098 < DECIPERE (Russell-Gebbett 1965, 64), *no la vol redre* 1080–1095 < REDDERE (Russell-Gebbett 1965, 75), *con omo debet eser a son sengore* 1035–1076 < *ESSERE, *qui tolre te-n buegan* 1035–1076 < TOLLERE (Russell-Gebbett 1965, 64),¹³ *o faré fer ad meus filium* ca. 1050 < FACERE (Russell-Gebbett 1965, 77), *et si non lo pod guarir* 1094–1110 < Germ. WARJAN (Russell-Gebbett 1965, 68).

4.4.2 Present indicative

Examples of the present indicative: *lo té de za ad riba de Scio* 1036–1079 < TENET (Russell-Gebbett 1965, 65); *e sson-o del fil* 1080–1095, *e sso rancuros de Guilelm Arnal* id. < SUM (Russell-Gebbett 1965, 75); *e si es om o femina* late 11th cent., *Guilelm Arnal es ab mi* 1080–1095 < EST (Russell-Gebbett 1965 71, 74); *chi ara i son* 1131 < SUNT (Russell-Gebbett 1965, 79); *rancur-me-n de la cavalleria* 1080–1095 < *RANCŪRŌ; *no la vol redre* (1080–1095) < VOLET (Russell-Gebbett 1965, 75); *qual nudriment i fa hom* (1085) < FACIT (Rasico 2006, 291); *que fan a Senta Maria* 12th cent. < FACIUNT (Russell-Gebbett 1965, 82); *et Ramon comes receb-le-n* mid-11th cent. < RECIPIT (Russell-Gebbett 1965, 68); *et y venez* 1036–1079 < VINCIT (Russell-Gebbett 1965, 65); *et fenex-la-li Pelet* 1117 < FĪNISCIT (Russell-Gebbett 1965, 69); *e aquel mor-i de fam* 12th cent. < *MORIT (Russell-Gebbett 1965, 82).

¹³ Regarding the context of this example, Russell-Gebbett (1965, 225) notes that *buegan* appears to derive from the subjunctive VOLEANT, although as he observes both the form of the verb (with *ue*, if derived from /ʷ/ followed by a yod) as well as the immediate context, do not support such an origin. However, a slightly different analysis, that which is proposed here, suggests **bulgan* (< *vulgan*), for OCat. *vulguen* < VOLEANT, due to scribal error or a false reading of the original 11th cent. manuscript.

4.4.3 Imperfect indicative

Examples: *res vos en passava o vos en frania* 1072–1079 < PASSĀBAT/FRANGIĒBAT (Russell-Gebbett 1965, 78); *o le-n frangien ad Bernardo* 1042–1117 < FRANGIĒBANT (Russell-Gebbett 1965, 67); *et si esbalczave lo castell* 1085 < *EXBALTEĀBAT (Rasico 2006, 291); *quar n'ere mes* 1080–1095 < ERAT (Russell-Gebbett 1965, 75); *si-s partia de mi* 1080–1095 < PARTIĒBAT; *com lo-m avia* (id.) < HABĒBAT; *no me-n jachia re* id. < Germ. *JEHHJAN (Russell-Gebbett 1965, 74); *et .I. somera que valie .III. solidos* 12th cent. < VALĒBAT (Rasico 2006, 291); *e com lo se-n-menave pres* id. < MINĀBAT; *ez-els tenien la casa en poder* id. < TENĒBAT (Russell-Gebbett 1965, 83).

4.4.4 Perfect indicative

The Old Catalan perfect preserves forms derived from both the Classical Latin strong and weak perfect paradigms. In the case of the former, the stress fell upon the stem of the verb in first and third persons singular; while in the weak perfect, the stress fell upon the ending in all paradigmatic forms (Badia i Margarit 1981a, 352–366; Moll 1991, 156–162; Coromines, 1974, 319–324).

Examples of both strong and weak perfects follow: *hon jo pris podstad* 1080–1095 < *PRĒNSĪ; *et que ag de suo pater* id. < HABUIT (Russell-Gebbett 1965, 74); *venit ad meo pater et sustrax-lo-li* late 11th cent. < SUBTRAXIT; *et no les li atès* id. < *ATTĒNSIT (Rasico 2015, 158); *e per zo car el no volie enar pres-lo pez collons* 12th cent. < *PRĒNSIT (Russell-Gebbett 1965, 83); *e agren-ne .III. besties* id. < HABUERUNT (Russell-Gebbett 1965, 84); *et Mir Arnall fer no-ls me volg* id. < VOLUIT; *et fed-le-n fer convenença* id. < FĒCIT; *et conveng-m'o* id. < CONVENUIT (Russell-Gebbett 1965, 74); *e mes-les en pret* id. < MISIT (Rasico 2015, 158); *trasc-ne .I. libre* id. < *TRASQUIT (Russell-Gebbett 1965, 82); *tolgren als omens d'Astafrancs* id. < *TOLLUERUNT; *e tengren-lo .III. dies pres en la tore* id. < TENUERUNT (Rasico 2006, 291); *.XL. uncias che manlevá de Berenger Bernard* 1094–1110 < MANU-LEVĀVIT (Russell-Gebbett 1965, 69); *que cavalgà Arnal Arnal super suo senior* 11th cent. < CABALLICĀVIT (Rasico 2015, 158); *et comanná-le-n Mir Arnal* 1080–1095 < COMMANDĀVIT; *che jo no-l li doné* id. < DONĀVĪ; *e rredí-l a Guilelm Arnall* id. < REDDĪVĪ (Russell-Gebbett 1965, 74); *e proclamà que-l me redés* late 11th cent. < PROCLAMĀVIT; *que chaçà meos conilos* id. < CAPTIĀVIT (Rasico 2015, 158); *e rééme-l .LXXX. solidos* 12th cent. < *REDIMIVIT; *menaren-lo a Cunquabela* id. < MINAVERUNT; *e trencaren la francea* id. < *TRENCAVERUNT; *preseren-se-n .III. milia* id. < PRĒNSĒRUNT (Rasico 2006, 291); *e de las folias que li dexist* 1080–1095 < LAXĪSTĪ (Russell-Gebbett 1965, 75).

4.4.5 Future indicative

Examples of the future, composed of the Vulgar Latin analytic combination of an infinitive followed by shortened forms of the present indicative of HABĒRE (later agglutinated as morphemes indicating person, number and tense to form new synthetic future forms) are: *go fideles vos ende seré* 1028–1047 < ESSERE-**HAIO* (Russell-Gebbett 1965, 63); *et potestate los en daré* 1035–1076 < DĀRE-**HAIO* (Russell-Gebbett 1965, 64); *si-o farà [...] aut o rechulirá de Remon Gonball* 1036–1079 < FACERE-**HAT*/RECOLLIGERE-**HAT* (Russell-Gebbett 1965, 65); *no llos vos farà* 1028–1047 < FACERE-**HAIO* ; *go a vos ende atenderé* id. < ATTENDERE-**HAIO* (Russell-Gebbett 1965, 63); *de esta ora ad avante non ti deceberé* 1035–1076, *de ista ora in antea nu dedebré Reimundus* 1039–1050 < DECIPERE-**HAIO* (Russell-Gebbett 1965, 64, 76); *no-t engannarei* late 11th cent. < **INGANNĀRE-HAIO* (Russell-Gebbett 1965, 71); *sí ti-o tenré et ti atendrè* 1043–1098 < TENERE-**HAIO*/ATTENDERE-**HAIO* (Russell-Gebbett 1965, 66).

4.4.6 Conditional indicative

Similar in formation to the future, the conditional paradigm (a future set in relation to a point in the past), is formed by the infinitive with agglutinated, reduced forms of the imperfect indicative of HABĒRE (Badia i Margarit 1981a, 372–375; Moll 1991, 167–168). Examples discovered in Old Catalan documents are: *que-l ne pagarie et non fecit* late 11th cent. < PACĀRE-**HĒBAT*; *que-ns en farie ço que fer* (id.) < FACERE-**HĒBAT* (Rasico 2015, 159); *no siria mos dons* 1080–1095 < **ESSERE-**HĒBAT (Russell-Gebbett 1965, 74);¹⁴ *lo don qui venria ad Raimundo comite* 1043–1098? < VENĪRE-**HĒBAT* (Russell-Gebbett 1965, 76); *fet-lo jurar manves que no se-n clamarie* 12th cent < CLAMĀRE-**HĒBAT* (Russell-Gebbett 1965, 83).

4.4.7 Perfect indicative paradigms

Analytical perfect paradigms were developed in Romance consisting of an auxiliary verb (in which person, number, tense and mood were expressed) followed by a perfect participle derived from the Latin passive participle. In Old Catalan the auxiliary *ésser* (~ *ser*) < **ESSERE* accompanied participles of intransitive verbs, while *haver* < HABĒRE was used with the participles of transitive verbs. In the case of the latter, the perfect

¹⁴ Russell-Gebbett (1965, 272) glosses *dons* as 'lord' < DOMINUS. However, given the context, *dons* would appear instead to represent Latin DAMNŌS 'harms', 'injuries'; cf. OCat. *don* (< **daun*) < DAMNU (MCat. *dany*), *escon* 'seat' < SCAMNU (Coromines ²1974, 262).

participle generally agreed in number and gender with the direct object (Badia i Margarit 1981a, 366–369; Moll 1991, 148; Coromines 1974, 325–326).

An example of the perfect infinitive is seen in *.LXXXIII. morabetins que degre aver pagads* 1190–1210 < (DĒBUERAT) HABĒRE PACĀTOS (Russell-Gebbett 1965, 85).

Examples of the present perfect indicative: *et ad jamdictam Elisabez jurad lur en e* 1039–1050 < IURĀTU *HAIO (Russell-Gebbett 1965, 76); *que an la casa de Sen Pere afolada e desfeita* 12th cent. < *HANT A-FULLĀTA/DISFACTA; *a la onor venuda e-n-pignorada* id. < HABET VINDŪTA/IMPIGNORĀTA; *les conamines que ssolien laurar ab la jova a tornades a sensum* id. < HABET TORNĀTAS (Russell-Gebbett 1965, 82); *coses que na Ramia e-n G. de Pons an preses* 12th cent. < HABENT PRĒNSAS (Rasico 2006, 291); *e no-n ha negu pagad* 1190–1210 < HABET PACĀTU (Russell-Gebbett 1965, 85).

Examples of the pluperfect indicative: *che Mir Arnall m'avía fait* 1080–1095 < HABĒBAT FACTU (Russell-Gebbett 1965, 74); *dret que avie fermat* 12th cent. < HABĒBAT FIRMĀTU; *e-ls esplets qe avien feit de .III. ans* (id.) < HABĒBANT FACTU (Rasico 2006, 291).

4.4.8 Subjunctive

Examples of the present subjunctive found in early Catalan texts are: *che chavalchar non puxa aut de aqua que passar non puxa* 1031–1035 < *POSSIAT (Moran/Rabella 2001, 50); *et hec omnia convenio tibi ched-t'o faça et t'o atena tot sine tuo engan* ca. 1088 < FACIAT/ATTENDAT (Russell-Gebbett 1965, 79); *tro li façan donar potestatem de illo castello* ca. 1043–1098 < FACIANT; *que sí le-n ajud Pere Mir* id. < ADIŪTET (Russell-Gebbett 1965, 76); *quantra totes omnes qui te-n tolran o qui tolre te-n buegan* 11th cent. < VOLEANT (Russell-Gebbett 1965, 64);¹⁵ *et convenit li Artall che li pledeg lo termen de Castilgon* 1094–1110, derived from *pled* (< *pleid*) < *PLACITU + *-e(i)g* < -IDIET (Russell-Gebbett 1965, 68); *ans que sis meses sien passats sie publicada* 12th cent. < *SIANT/*SIAT (Russell-Gebbett 1965, 80); *che ille no la li rancur* 1117 < *RANCŪRET (Russell-Gebbett 1965, 69); *que s'endrez tot aquest feit* 12th cent. < *INDIRĒCTIET (Russell-Gebbett 1965, 83); *la tersa part de la tascha de tota re que-n escha* 1189 < EXIAT; *e de cascú mas de sas pagesias on bèstias aja* id. < HABEAT (Moran/Rabella 2001, 84).

The past subjunctive appears in the following examples: *tro Bonfilg le-n agéss emendat lo don* 1043–1098? < HABUISSET (Russell-Gebbett 1965, 76); *contra tots homines et contra totas feminas qui tolre-ls-te volgesen o-ls te tolgesen* 1048 < VOLUISSENT/*TOLLUISSENT (Russell-Gebbett 1965, 77); *et si Guilelm Arnal me facia tal cosa*

¹⁵ Moran/Rabella (2001, 54), who follow closely Russell-Gebbett's edition, suggest 1047–1098 as the possible period of composition of this feudal oath. On the subjunctive verb form *buegan*, which is transcribed by both Russell-Gebbett and Moran/Rabella, see 4.4.1 above.

que dreçar no-m volgés ho no pogés 1080–1095 < VOLUISSET/POTUISSET; che Mir Arnall me romasés id. < *REMANSISSET; che no-m estacás dret e che no-l me fedés sens el id., derived from *estaca* (< Germ. *STAKKA)/FECISSET (Russell-Gebbett 1965, 74); et proclamà que-l me redés late 11th cent. < REDDISSET (Rasico 2015, 159).

5 Bibliography

- Alart, Julien B. (1877), *Études historiques sur quelques particularités de la langue catalane*, Revue des Langues Romanes 12, 109–132.
- Arasa i Gil, Ferran (1994–1995), *Aproximació a l'estudi del canvi lingüístic en el període ibero-romà (segles II–I a.C.)*, Arse. Boletín del Centro Arqueológico Saguntino 28–29, 83–107.
- Badia i Margarit, Antoni M. (1981a), *Gramàtica històrica catalana*, València, Tres i Quatre. Spanish edition: (1951), *Gramàtica històrica catalana*, Barcelona, Noguer.
- Badia i Margarit, Antoni M. (1981b), *La formació de la llengua catalana. Assaig d'interpretació històrica*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Baraut, Cebrià (1978), *Les actes de consagracions d'esglésies del bisbat d'Urgell (segles IX–XII)*, Urgellia 1, 11–182.
- Baraut, Cebrià (1984–1985), *La data de consagració de la catedral carolíngia de la Seu d'Urgell*, Urgellia 7, 515–529.
- Bastardas, Joan (1995), *La llengua catalana mil anys enrere*, Barcelona, Curial.
- Batlle, Mar et al. (2016), *Gramàtica històrica de la llengua catalana*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Bisson, Thomas N. (1988), *Història de la Corona d'Aragó a l'Edat Mitjana*, Catalan translation by Mariona Vilalta, Barcelona, Crítica. Original English edition: (1986), *The Medieval Crown of Aragon. A Short History*, Oxford, Oxford University Press.
- Coromines, Joan (1954), *El que s'ha de saber de la llengua catalana*, Palma de Mallorca, Moll.
- Coromines, Joan (1965–1970), *Estudis de toponímia catalana*, 2 vol., Barcelona, Barcino.
- Coromines, Joan (?1974), *Lleures i converses d'un filòleg*, Barcelona, Club Editor.
- Coromines, Joan (1976–1977), *Entre dos llenguatges*, 3 vol., Barcelona, Curial.
- DCVB = Alcover, Antoni M./Moll, Francesc de B. (1927–1962), *Diccionari català-valencià-balear*, 10 vol., Palma de Mallorca, Moll. Reprint 1980.
- DECat = Coromines, Joan (1980–2001), *Diccionari etimològic i complementari de la llengua catalana*, 10 vol., Barcelona, Curial/La Caixa.
- Ferrando Francés, Antoni/Nicolás Amorós, Miquel (2011), *Història de la llengua catalana*, Barcelona, UOC.
- Gili Gaya, Samuel (1955), *Notas sobre el mozárabe de la Baja Cataluña*, in: Antonio Badía Margarit/Antonio Griera/Federico Udina Martorell (edd.), *VII Congreso Internacional de Lingüística Románica*, Barcelona, Abadía de San Cugat del Vallés, vol. 1, 483–492.
- Grandgent, Charles H. (1905), *An Outline of the Phonology and Morphology of Old Provençal*, Boston, D.C. Heath & Co.
- Guiter, Henri (1960), *Phonétique évolutive et toponymie historique*, Revue des Langues Romanes 74, 1–53.
- Gulsoy, Joseph (1993), *Estudis de gramàtica històrica*, València/Barcelona, Institut Universitari de Filologia Valenciana/Publicacions de l'Abadia de Montserrat.
- Lleal, Coloma (1990), *La formación de las lenguas romances peninsulares*, Barcelona, Barcanova.
- Miquel Rosell, Francisco (ed.) (1945), *Liber Feudorum Maior: cartulario real que se conserva en el Archivo de la Corona de Aragón*, 2 vol., Barcelona, Consejo Superior de Investigaciones Científicas.

- Miret i Sans, Joaquim (1908), *Documents en langue catalane (haute vallée du Sègre, XI–XII siècles)*, *Revue Hispanique* 19, 6–19.
- Miret i Sans, Joaquim (1912), *Aplech de documents dels segles XI y XII per a l'estudi de la llengua catalana*, *Boletín de la Real Academia de Buenas Letras de Barcelona* 6, 348–357, 381–395.
- Moll, Francesc de B. (1991), *Gramàtica històrica catalana*, València, Universitat de València. Spanish edition: (1952), *Gramática histórica catalana*, Madrid, Gredos.
- Moran i Ocerinjauregui, Josep (1994), *Treballs de lingüística històrica catalana*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Moran i Ocerinjauregui, Josep (1995), *Estudis d'onomàstica catalana*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Moran i Ocerinjauregui, Josep (2006), *El procés de creació del català escrit*, in: Philip D. Rasico (ed.), *El català antic*, Girona, Universitat de Girona/CCG Edicions, 7–48.
- Moran, Josep/Rabella, Joan Anton (edd.) (2001), *Primers textos de la llengua catalana*, Barcelona, Proa.
- Moran, Josep/Rabella, Joan Anton (2015), *The Process of Scripturising Catalan*, in: Flocel Sabaté/Luís Adão da Fonseca (edd.), *Catalonia and Portugal: The Iberian Peninsula from the Periphery*, Bern, Lang, 117–130.
- Mundó, Anscari M. (1974), *Domains and Rights of Sant Pere de Vilamajor (Catalonia): A Polyptych of c.950 and c.1060*, *Speculum* 49, 238–257.
- OnCat = Coromines, Joan (1989–1997), *Onomasticon Cataloniae. Els noms de lloc i noms de persona de totes les terres de llengua catalana*, 8 vol., Barcelona, Curial/La Caixa.
- Pujol, Pere (1913), *Documents en vulgar dels segles XI, XII & XIII procedents del bisbat de la Seu d'Urgell*, Barcelona, Biblioteca Filològica de l'Institut de la Llengua Catalana.
- Pujol, Pere (1917), *L'acte de consagració i dotació de la catedral d'Urgell, de l'any 819 o 839*, *Estudis Romànics (Llengua i Literatura)* 2, 92–115.
- Rabella i Ribas, Joan Anton (2012), *Orality i escriptura: la llengua catalana a l'Edat Mitjana*, in: Juan Pedro Sánchez Méndez (ed.), *Oralidad y escritura en la Edad Media hispánica*, Valencia, Universitat de Neuchâtel/Tirant Humanidades, 53–85.
- Rasico, Philip D. (1982), *Estudis sobre la fonologia del català preliterari*, Barcelona, Publicacions de l'Abadia de Montserrat/Curial.
- Rasico, Philip D. (1993), *Estudis i documents de lingüística històrica catalana*, Barcelona, Curial.
- Rasico, Philip D. (2006), *El català antic*, Girona, Universitat de Girona/CCG Edicions.
- Rasico, Philip D. (2015), *Cries of Abuse and Injustice in Early Catalan: Notes on the Language of the "Rancures", "Clams" and "Querimònies" (11th and 12th Centuries)*, in: Flocel Sabaté/Luís Adão da Fonseca (edd.), *Catalonia and Portugal: The Iberian Peninsula from the Periphery*, Bern, Lang, 131–164.
- Recasens i Vives, Daniel (2017), *Fonètica històrica catalana*, Barcelona, Institut d'Estudis Catalans.
- Russell-Gebbett, Paul (1965), *Mediaeval Catalan Linguistic Texts*, Oxford, Dolphin.
- Sanchis Guarner, Manuel (1950), *Introducción a la historia lingüística de Valencia*, Valencia, Institución Alfonso el Magnánimo.
- Sanchis Guarner, Manuel (²1961), *Els parlars romànics de València i Mallorca anteriors a la Reconquesta*, Valencia, Institución Alfonso el Magnánimo.
- Sanchis Guarner, Manuel (⁴1972), *La llengua dels valencians*, Sueca, l'Estel.
- Sanmartí, Joan (2009), *From the Archaic States to Romanization: A Historical and Evolutionary Perspective on the Iberians*, *Catalan Historical Review* 2, 9–32.
- Soldevila, Ferran (²1963), *Història de Catalunya*, Barcelona, Alpha.
- Turull, Albert (2007), *La toponímia de les comarques de ponent. Un assaig d'interpretació tipològica*, Barcelona, Institut d'Estudis Catalans.

- Udina Martorell, Federico (ed.) (1951), *El Archivo Condal de Barcelona en los siglos IX–X*, Barcelona, Consejo Superior de Investigaciones Científicas.
- Veny, Joan (1991), *Els parlars catalans*, Palma de Mallorca, Moll.

Antoni Ferrando

11 The Growth and Expansion of Catalan (1213–1516)

Abstract: During the period of three centuries that extends from 1213, the year of King Peter the Catholic's defeat at Muret, marking the end of the Catalan political presence in Occitania, to 1516 when, with the death of King Ferdinand II, the Crown of Aragon was integrated into a single Hispanic monarchy, the Catalan linguistic community underwent a process of vigorous territorial, political and cultural expansion. Over the course of this time, the Catalan language became clearly separated from Occitan; it came to share the political space with the Aragonese language; it took its place as the primary language of the royal domains; it became the oral and written heritage of all the strata of Catalan-speaking society; it formed an accepted Chancery *koine* for all spheres of the public administration and for all cultured literary registers; it spread its influence to Italy, and it gave rise to notable works in all literary genres, from the prose of Ramon Llull and the four *Great Chronicles* to the verse and narratives of the outstanding writers of the 15th century – the Golden Age of Catalan literature –, including Ausiàs March, Joanot Martorell and Joan Roís de Corella.

Keywords: Catalan language, Catalan scripta, territorial expansion, Chancery language, Golden Age of Catalan literature

1 Historical and territorial framework

At the start of the 13th century, the Catalan language was limited to the historical territory of Catalonia, which comprised both the land situated to the south of the Pyrenees (the future Principality of Catalonia and part of the Catalan-speaking strip of the Kingdom of Aragon) and the four North Pyrenean counties (Rosselló, Cerdanya, Conflent and Vallespir). By that time the Counts of Barcelona had succeeded in dominating all the Catalan counties except Empúries, Urgell and Pallars Sobirà, which would come under the royal crown in 1402, 1413 and 1487, respectively. The Counts of Barcelona had also succeeded in making a substantial number of Occitan counties their vassals. It may be recalled here that, thanks to the marriage between Ramon Berenguer IV, Count of Barcelona, and Peronella, Queen of Aragon, the Catalan-Aragonese confederation had been formed in 1137.

The French intervention in Occitania, under the pretext of defending Catholic orthodoxy against the Albigensian heretics, led to the defeat of King Peter the Catholic at Muret (1213) and, consequently, to the loss of the Catalan possessions in Occitania, except for the domain of Montpellier (which was maintained until the

<https://doi.org/10.1515/9783110450408-018>

mid-14th century). This loss was more than compensated by the incorporation of the Balearic Islands (1229, 1287), of the Pityusic Islands (1235) and of the lands of Sharq al-Andalus (1232–1245) into the Crown of Aragon. Thanks to this military and colonising action of King James I, almost the entire present-day Catalan linguistic area came to be formed. James I's grandson, James II, completed it with the incorporation of the governorate of Oriola (1304). The political union between Aragon and Catalonia entailed not only the consolidation and expansion of the Catalan language but also the preservation of the Aragonese language, which had been highly exposed to the impact of Castilian. The language which benefited most from this union, however, was Catalan since the monarchs sought to ensure its dominance in the recently conquered kingdoms. Nevertheless, the common monarch of the Catalan-Aragonese Crown as a whole received the title of King of Aragon as this was the highest rank of nobility.

King James I created the Kingdom of Majorca with the Balearic Islands, the four North Pyrenean counties and the domain of Montpellier (1276–1343). This kingdom was reintegrated into the Crown of Aragon, however, by the military action of Peter the Ceremonious against James III of Majorca. On the other hand, James I turned the lands conquered in Sharq al-Andalus into the Kingdom of València (1238), which lasted until the decree of annexation of València by the Kingdom of Castile (1707) and shared the same monarchs with Catalonia and Aragon until the decrees of Nueva Planta (1707–1716).

The Crown of Aragon was ruled by the House of Barcelona until 1410, when King Martin the Humane died without issue. All three kingdoms agreed in the Compromise of Caspe (1412) to grant the Crown of Aragon to the Castilian prince Ferdinand of Antequera, a grandson of the Catalan-Aragonese king Peter the Ceremonious, thus ushering in the Castilian House of Trastámara (1412–1516).

The dynastic union of the Principality of Catalonia – a designation which appears on record from the 13th century – with the Kingdom of Aragon came to add cultural and political complexity to the Crown of Aragon as well as allowing it to affirm itself as a political and economic power. It was a strategic union of mutual convenience, which helped not only to curb the expansionist and annexationist ambitions of France and of Castile but also to lay the foundations for the creation of an important Mediterranean empire. This mostly commercial empire, competing with other maritime powers such as Genoa and Venice, was already a reality in the 13th century. The Crown of Aragon reached its territorial and political plenitude in the 14th century with the incorporation of Sardinia (1323) and, above all, in the 15th century with the incorporation of the Kingdom of Sicily (1409) – which had been governed by the House of Barcelona since 1282 – and of the Kingdom of Naples (1442–1458), where King Alfons the Magnanimous established the court of the entire Crown of Aragon. The military and colonial expansion came to comprise briefly the greater part of the Kingdom of Murcia (1266–1304) and the duchies of Athens and Neopatra (1379–1390). There was little presence of any Catalan population in these Mediterranean

territories. The city of l'Alguer in Sardinia, which was resettled in 1354, is the only vestige of Catalan colonisation on the island.

There were four distinct situations within the political-administrative structure of the Crown of Aragon: a) the Catalan-speaking territorial strip in the Kingdom of Aragon, which preserved the administrative uses of Catalan throughout this whole Late Medieval period; b) the dozen mainly Aragonese-speaking towns in the western area of the Kingdom of València, where Catalan had a limited presence in the administrative sphere; c) the Catalan-speaking domain of Andorra, which succeeded in preserving its independence against the neighbouring powers thanks to the condominium – the *paréage* agreement of 1278 – of the bishops of Seu d'Urgell, a Catalan diocese, and of the Counts of Foix, vassals of the Kings of France; and d) the Occitan-speaking domain of the Valley of Aran, which initially formed part of the Kingdom of Aragon and was finally incorporated into Catalonia in 1313.

In the 15th century the Aragonese Trastámaras focused their international interests on two fronts: Castile, as a result of their family ties, and Italy, as a continuation of a political path of the House of Barcelona which started in the 13th century with the marriage of prince Peter of Aragon, the future King Peter the Great, and Constance of Hohenstaufen, the heiress to Sicily and Naples. The attention to Castilian affairs and the ties of the Kingdom of Navarre (1441–1479) to the Crown of Aragon strengthened the presence of Castilian at the Catalan-Aragonese court. On the other hand, the presence of Catalan-speaking ecclesiastics (and, consequently, of the Catalan language) at the papal court in Rome was so notable that two of them, the cardinals Alfons de Borja and Roderic de Borja, became popes under the names of Callixtus III (1455–1458) and Alexander VI (1492–1503), respectively.

In the medium term, however, the Trastámaras' Castilian interests came to prevail: the dynastic union between the Crown of Aragon and the Kingdom of Castile took place in 1479 when the married couple formed by Ferdinand II of Aragon and Isabella I of Castile (who had already been a queen since 1474) became their monarchs. Although this union maintained the legal sovereignty of both states, it benefited de facto Castile since it came to incorporate into Castile's domains the Muslim Kingdom of Granada (1492), the southern part of the Kingdom of Navarre and the greater part of Central and South America (as from 1492). The Crown of Aragon, on the other hand, only succeeded in reincorporating the Kingdom of Naples (1503). According to the marriage covenant of the Catholic Monarchs, Ferdinand II was obliged to reside in Castile, so the court common to the two states became de facto a Castilian court. The year 1479 marked, therefore, the start of a process of political and cultural subordination of the Crown of Aragon with respect to Castile, which was economically, demographically and territorially a much more powerful kingdom.

The period of interest to us here ends with the death of Ferdinand II (1516), the last king to reign exclusively over the Crown of Aragon. Subsequently, Castile and the Crown of Aragon would share a single monarch, Charles I. Indeed, it was then that the

diarchy of the Catholic Monarchs would really become the Hispanic Habsburg Monarchy (Gabriel 1994–1999; Riquer 1995–2000).

2 Catalan, a Romance language: *Scripta* and linguistic evolution

Medieval Catalan shared with Occitan the adverb of affirmation *oc*, a linguistic trait which separates them from the neighbouring Romance languages. This coincidence makes their common origin clearly evident. The two languages were originally so close that, in the field of Romance philology, the criterion of considering Catalan to be a distinct language from Occitan was not established until the early 20th century (Sanchis Guarner 1980; Colón 1993). In fact, Catalan only came to clearly distance itself from the rest of the Occitan linguistic diasystem as from the 13th century, as a result of the political decoupling of the Crown of Aragon and Occitania. The political and cultural subordination of Occitania to the Kings of France as from the 13th century and Catalonia's legal emancipation from the Kings of France by the Treaty of Corbeil (1258) contributed to the consolidation of Catalan as a distinct language from Occitan. It was then that Catalan began to adopt its own *scripta*, distinct from the Occitan, such as the preference for the digraphs *ny* (*Espanya*) and *ll* (*fill*) as opposed to *nh* (*Espanha*) and *lh* (*filh*) in Occitan, respectively. Such choices evidence Catalan's will of emancipation from Occitan.

The territorial expansion of Catalan to the Balearic Islands and to València also entailed the extension of its two constitutive dialects: there where the colonists from the area between Barcelona and Rosselló predominated, specifically in the Balearic Islands, Eastern Catalan prevailed; and there where the colonists from the “Terra Ferma” area (Urgell, Ribagorça, Lleida, etc.) predominated, specifically in the Kingdom of València, Western Catalan prevailed, even though there were substantial numbers of Eastern Catalans and Aragonese in these areas too. The demographic weight of the Western Catalans and the linguistic levelling between these colonising components favoured the triumph of the Western form of Catalan in the Kingdom of València (Ferrando 1989). It had been traditionally considered that the Western character of Valencian was due to reasons of substrate (Sanchis Guarner 1980; Badia i Margarit 1981). In this way the two “constitutive” dialects of Old Catalonia, which had spread to New Catalonia (the territories conquered as from 1148), extended as “consecutive” dialects to the lands taken from the Eastern Andalusia. The medieval Balearic forms of speech reflected the traits of the northeastern forms of speech of Catalonia (such as the resulting tonic neutral *e* derived from the close E of Vulgar Latin, the tendency of confusion between atonic *a* and *e*, or the *salat* article: use of *es*, *sa* or *so* instead of *el*, *la* and *lo*). On the other hand, the Catalan of the Valencian lands preserved the most prominent features of Northwestern Catalan (the tonic close *e*

derived from the close E of Vulgar Latin, the distinction between atonic *a* and *e*, the tendency towards inchoatives in *-isc*, etc.). L'Alguer still maintains today some linguistic traits of the 14th-century continental Catalan (such as the desinence *-au* of the second person plural of the present indicative instead of *-eu*). (78.1 Dialects).

Just as all languages, medieval Catalan evolved. If one compares its changes with those of Occitan, the changes of Catalan reached much deeper. For example, while Catalan already began in the 13th century to vocalise the ending *-ts* of the second person plural of certain tenses (whereby such forms as *portats*, *batets*, *partits* became *portau*, *bateu*, *partiu*), Occitan did not cease to maintain *-tz* (*portatz*, *batetz*, *partitz*). The changes undergone by Catalan had a great deal to do not only with its decoupling from the Occitan world, but also with its constant contact with Aragonese, the growing weight of the Western dialect – lexically more closely allied to Aragonese and Castilian –, and, with respect to the literary language, the impact of Castilian as from the 15th century, favoured by its status as a court language, together with Catalan, and as the personal language of the Aragonese Trastámaras. The influence of Aragonese was quite notable in the colloquial Catalan of València, which generally adapted loan words to the structure of Catalan (*corder*, *a boqueta nit*, etc.), but not always (*ganado*, *bonico*, etc.). Some of the Aragonese loans adapted or semiadapted to Catalan (*fardatxo*, *foia*, *brullo*, *tomello*, etc.) have been wrongly interpreted as Mozarabic words (Ferrando/Nicolás 2011). All this contributed to the start of the so called lexical hispanization of Catalan. This phenomenon explains the literary preference for certain Ibero-Romance forms such as *castigar*, *fartar*, *deixar*, *casar*, *baixar*, etc., instead of the respective Gallo-romance forms: *punir*, *sadollar*, *lleixar*, *maridar*, *devallar*, etc. (Colón 1993). The whole of Catalan did not undergo the same lexical changes. The lesser impact of exogenous factors in the Balearic dialects of Catalan explains the conservatism of their vocabulary.

In the last half of the 15th century, the predominance acquired by the bourgeois classes in Valencia and Catalonia manifested itself in the written language through the emergence of a set of linguistic traits which were already quite consolidated orally, such as the generalisation of the desinence *-e* of the first person singular of the present indicative of the verbs of the first conjugation (*porte*, *mire*, etc.), the generalisation of the plurals in *-os* (*peixes* > *peixos*) and of the previously mentioned vocalised verbal forms (*amau*, *teníeu*, *partiu*, etc.), the replacement of *ésser* by *haver* as an auxiliary in the compound tenses of the pronominal verbs (*só restat* > *he restat*), and the tendency to prefer the order IO + DO in the binary combinations of weak pronouns (*la'm porta* > *me la porta*). In Eastern Catalan, the confusion between atonic *a* and *e* became generalised at the same time, and the closing of the atonic *o* into *u* (*portava* > [pur'taβə]) began, which were oral changes that good written language tended to avoid. On the other hand, the Catalan of the Balearic Islands preserved many of its archaic traits beyond the 15th century (Coromines 1971).

The contact with other languages also affected the physiognomy of Catalan. Despite the more or less generalised linguistic isolation between Catalan speakers and

Arabic speakers, the persistence of Arabic in New Catalonia, the Kingdom of València and the Balearic and Pityusic Islands explains the transfusion of many Arabic loans into Catalan and particularly into the Catalan of the Ebre Valley and of València (*alfarrasar*, *carxofa*, *cotó*, *séquia*, etc.) (Barceló 2011). Already in the 15th century, with the enthronement of the Trastámaras, the prestige of Castilian among the native nobility favoured the penetration of courtly Castilian words (*aposeno*, *despedir*, etc.). Likewise, the contacts with Italy were reflected in the adoption or adaptation of such words as *infanteria*, *escaramussa*, *mercant*, etc. On the other hand, Catalan contributed to the *sabir* or *lingua franca* of the merchants of the Mediterranean with such words as *tafurea*, and it influenced Sardinian, Sicilian and Neapolitan.

3 Linguistic awareness: The onomastic question

Catalonia's political decoupling from the Occitan territories did not entail a complete cultural break between Catalans and Occitans. Catalan continued to share with the rest of Occitan a brilliant troubadour literature up to and including the 13th century, which persisted as a courtly poetic language in the Crown of Aragon until the first quarter of the 15th century. The decoupling process, however, favoured the awareness of a linguistic distancing, which explains why the Occitan texts began to be translated into Catalan. The *Homilies d'Organyà* already came to be translated in the early 13th century. The *Vides de sants roselloneses* (last half of the 13th century), of a more cultivated intention, also seem to be a translation from Occitan. The *Breviari d'amor*, the vast encyclopaedia in verse of Matfrè Ermengaud de Besiers, was translated in the 14th century. All this shows that Occitan, which was usually called *llemosí* or *provençal* in the Crown of Aragon, already came to be perceived in the 13th century as a language distinct from Catalan (Nadal/Prats 1982).

The most ancient name for Catalan was *catalanesc*, which was repeatedly used in the *Crònica* of Ramon Muntaner, but the syntagm *llengua catalana* is on record from the early decades of the 14th century. The political-administrative fragmentation of the Crown of Aragon explains why, from the 14th century, the demonyms of the new Catalan-speaking territories, *valencià* and *mallorquí*, were also applied to the Catalan language of those territories. These local or regional names are similar to those which may be found in other languages, such as Toulousain, Florentine, Pisan, Toledan, etc. In general, the particularist onomastic uses of Catalan were limited to the internal sphere of the respective political communities and they were compatible with the recognition of the territorial scope of the Catalan linguistic community and, consequently, with the name *català*. It should be pointed out, however, that the particularist names did not reflect diatopic variation, but rather the awareness of belonging to a specific political-administrative entity.

Printing, which was introduced in Valencia and Barcelona around 1472–1475, helped to consolidate the onomastic disaggregation of the language since books

tended to cover and reflect the local demands and, consequently, the particularist names. As a result of this, Catalans began to use the name *català* exclusively with respect to the language of Catalonia, and Valencians abandoned the earlier onomastic oscillations – *valencià* in internal documents and *català* in international relations – to opt exclusively for the name *valencià*. Above all, however, the integration of the Crown of Aragon in the monarchy of the Catholic Monarchs (1479) was the determining factor in the consolidation of the language's particularist names (Ferrando 1980).

4 Unity and diversity of the Catalan language: The Royal Chancery

Despite the dialectal bipartition and the onomastic diversity, medieval Catalan was a notably unitary language both in objective terms and in the perception of its speakers. The influence of the Royal Chancery had much to do with the unity of the written language, especially as from the time of its restructuring, which was ordered by James II (1319), and of the *Ordinacions de la Casa reial* (1344), of Peter the Ceremonious. The municipal scriiveneries of the foremost cities of the Crown of Aragon, particularly including that of the city of València, contributed decisively to the consolidation of the Chancery model. It should be kept in mind that Catalan shared with Latin and Aragonese the status of Chancery language. The good Latin training of the Chancery officials, the diversity of their geographical provenances and the initial centralising role of Barcelona were decisive factors in the formation of a cultured and supradialectal Chancery register. The language which emanated from the Royal Chancery was the most effective vehicle for ensuring the linguistic cohesion of the Crown of Aragon since, owing to its function as an administrative body of the sovereigns, it was able to impose some unifying linguistic patterns. Thanks to the prestige of this model, all the users of the cultured written language conformed to it (Ferrando 2016; 2018).

Beneath this unitary cultured language, however, numerous testimonies have been preserved of the diastratic and diatopic variation of Catalan. Didactic prose and legal and administrative documentation are an invaluable source for our knowledge of linguistic variation. Works such as *Vides de sants rosselloneses* (13th century); the anonymous translation of the *Diàlegs de Sant Gregori* (c. 1300), of northeastern origin; the *Regiment de preservació de pestilència* (1349), by Jaume d'Agramont, from Lleida; the version of the *Questa del Sant Graal* (1380), of Majorcan provenance; the *Ordenances dels corredors de llotja de Barcelona* (1271), and the *Llibres de Cort* of Montuïri, in Majorca (1357–1360 and 1415–1420), to mention just a few of the best-known texts, provide us with a wealth of information on the incipient phonetic and lexical colouring that each geographical variety was in the process of acquiring. Private letters and utilitarian texts are the documents which best show the contrast between the Chancery model and the local linguistic uses (Veny 1993). The *Regles d'esquivar vocables o*

mots grossers o pagesívols (c. 1492), attributed to Bernat Fenollar and Jeroni Pau, are a highly useful source for learning about the process of lexical regionalisation of Catalan towards the end of the 15th century.

The Chancery language lost influence as a result of the political, social and cultural changes in the closing decades of the 15th century (the civil war in Catalonia between 1462 and 1472, the dynastic union and the transfer of the royal court to Castile in 1479, the replacement in 1494 of the Royal Chancery by a Council of the Kingdoms of the Crown of Aragon – a sort of ministry for the affairs of this Crown –, and the inclination of Ferdinand II to use Castilian). The new court language was Castilian. Even so, printing helped to ensure the modernisation, regularisation and Latinisation of Catalan and to carry out in that way part of the uniformising functions which the Royal Chancery had previously performed (Ferrando 2000).

5 Social and institutional uses

In the Late Medieval centuries, Catalan was the most widespread instrument of communication of all the social strata of Catalonia, Valencia and the Balearic Islands, and it was used in all oral and written registers. Not only a large part of the administrative texts but also account books and books on cookery, astronomy, medicine, history and other subjects were written in Catalan (Vernet/Parés 2005). Already in the 13th century, Catalan displaced Latin as the language of justice and for numerous administrative uses in the municipal and royal spheres. Proof of this is the fact that, shortly after the mid-13th century, the first known manuscript of the Catalan version of the *Usatges* of Barcelona, the *Furs* of València, the *Costums* of Tortosa and the original nucleus of the *Llibre del Consolat de Mar*, which were texts of legal-administrative character, followed in succession. Needless to say, Latin continued to be used for the most prestigious purposes in the field of the civil and ecclesiastical administrations and in that of the most speculative sciences, such as theology and philosophy. But even so, laymen such as Ramon Llull, Arnau de Vilanova and Francesc de Pertusa used Catalan to write treatises on theology, a subject which was usually reserved for Latin. Likewise, numerous treatises on medicine in Arabic were translated into Catalan, and the translation of the Bible attributed to Friar Bonifaci Ferrer (1478) was the first printed Romance version of the Holy Scriptures.

Catalan was the language of approximately three-fourths of the non-Arabic-speaking population of the Crown of Aragon. This demolinguistic predominance and the fact that Catalan was the usual language of the royal house of Barcelona made it the vehicle of symbolic identification of the overall set of the royal domains. The Valencian Mudejars, who formed compact communities with recognised linguistic rights, preserved Andalusí Arabic as their usual language and their language of culture. On the other hand, the Mudejars of the Ebre Valley, who were demographically weak, and those of Majorca, who held the status of slaves, finally came to adopt

Catalan after a period of bilingualism. The Jewish communities, which were quite important in the 13th and 14th centuries, expressed themselves in Catalan, using Hebrew only for worship and for theological studies. Additionally, Catalonia's commercial relations ensured the presence of Catalan in dozens of port cities of the Mediterranean.

Until the 15th century, the use of Catalan and of Aragonese in the Royal Chancery did not have the weight shown by Castilian in the Chancery of the Kingdom of Castile. In Castile, Castilian overwhelmingly displaced Latin as the language of the royal administration as from the time of Alfonso X the Wise (1252–1284). In contrast, Latin maintained a very notable presence in the royal administration in the Crown of Aragon until the 15th century. This difference may be considered a consequence of the plurilingualism of a crown which, in the mid-15th century, came to be formed by three Hispanic and three Italic kingdoms. On the other hand, Catalan had a certain use as a Chancery and court language in Sicily, Naples and, above all, Sardinia, and it became one of the most widely used languages in the ecclesiastical courts of Rome in the last half of the 15th century and the opening years of the 16th. Together with Latin, Catalan was the language of the international relations of the Crown, except in dealings with the other Hispanic kingdoms and with the western countries of the Maghreb, in which Aragonese was preferred since it was more similar to Castilian.

The plurilingualism and the “confederal” structure of the Crown of Aragon help to explain the institutional respect for linguistic diversity. The languages of the domains of the Kings of Aragon (Latin, Catalan, Aragonese, Occitan, Arabic, Sardinian, Sicilian, Neapolitan) had a well-defined weight and distribution of communicative functions in each territory. Sometimes, however, the Royal Chancery and the local administrations respected personal linguistic uses.

The enthronement of the Trastámaras did not initially alter the sociolinguistic status of Catalan. The new kings, who all had Castilian as their mother tongue, continued to use it in the family and court settings but, subject to the laws and institutions of the various kingdoms of the Crown of Aragon, they respected and also used the Catalan language. This is why the private use of the monarchs' language hardly had any social or literary effects at all.

The change of linguistic paradigm took place under Ferdinand II. The high nobility, related through marriage to their Castilian counterparts or linked to the affairs of the monarchy, was the social sector which adapted most quickly to the new linguistic and cultural situation. The preference for Castilian affected not only these sectors of autochthonous nobility but also the officials in the service of the Council of Aragon, who had to follow the king on his travels. Catalan also became a minority language in international relations. With Ferdinand II, the initial courtly plurilingualism tended to quickly become a Castilian monolingualism. Moreover, the traditional antagonism of the subjects of the Crown of Aragon to everything Castilian soon turned to admiration, and knowing the king's language became an indispensable means of political and social advancement. Nevertheless, it should be said that the linguistic

behaviour of these social sectors had a greater effect on the public uses and the social prestige of Catalan than on the private uses and the vitality of the language in quantitative terms (Cahner 1980). In any case, Catalan went from a situation of social and territorial plenitude in the 15th century to one of political subordination, of decline in social prestige and reduction of use in high domains in the early 16th century.

6 Literary uses

The economic and demographic development of the Western Europe and the growing importance of the role of the “bourgeoisie” ensured the triumph of Catalan as a majority vehicle of cultural expression already in the 13th century. This opened the way for the secularisation of literature, a process led to a large extent by laymen. In this respect, a decisive role was played by the Majorcan Ramon Llull (c. 1232–1316), which was similar to that played by Alfonso X the Wise for Castilian. In Llull’s hands, Catalan achieved a considerable degree of syntactic maturity and an unquestionable expressive modernity (Badia i Margarit 2004). It is almost a commonplace to consider him the creator of Catalan literary prose. In fact, he perfected an already consolidated literary language, which may be observed in the prose of Arnau de Vilanova and in the two oldest Great Chronicles, the *Llibre dels feits del rei en Jaume* and the *Crònica del rei en Pere* (or Chronicle of Desclot). Under the pen of Llull and of Vilanova, Catalan also became a language of theology, and in that of Ramon Muntaner it became an effective narrative medium of political propaganda. Courtly poetry in the troubadours’ language was the only literary genre that was relatively impermeable to Catalan – but it was written in a conventional Occitan that exuded the native Catalan of its authors. Between the 12th and the beginning of the 15th centuries, the courtly Catalan poets such as Guillem de Berguedà, Ramon de Besalú, Cerverí de Girona, Jofré de Foixà, the brothers Jaume and Pere March, Gilabert de Pròixida or Andreu Febrer used the Occitan of the troubadours. Some members of the royal family, such as the kings Alfons the Troubadour, Peter the Great and Peter the Cerimonious, cultivated the language as well. However, from the third decade of the 15th century onwards, Ausiàs March and the majority of the poets of the court of Alfons the Magnanimous firmly opted for Catalan.

Standing out in the last half of the 14th century and the opening decades of the 15th are the *Crònica* of Peter the Ceremonious, conceived as a work of an exacting historian; the meticulous didactic prose of Francesc Eiximenis and the fiery words of Vicent Ferrer, both of them closely adhering to the medieval models and values, and the mordant satires of Anselm Turmeda, who better than any other writer expressed the true dimension of the spiritual crisis of the times of the Western Schism. With Peter the Ceremonious and his sons John I and Martin the Humane, the Royal Chancery also became a focus of cultural renewal, assimilating the new lay style of the pre-humanists, notable among whom were Bernat Metge and Antoni Canals,

sensitive to the Petrarchan model. At the same time, an intense translating activity was taking place, not only from Latin and Arabic and in some cases Greek, but also from French and from Occitan. The translations of scientific literature favoured the vernacularisation of science. And for its part, a short chivalric narrative set in the Ottoman Empire, *Jacob Xalabín*, from the early 15th century, reflects quite well the Mediterranean dimension of the Catalan language.

When the 15th century was in full swing, literary works of great value were produced which well justify this period's designation as the Golden Age of Catalan literature. These works notably include the poetry of Ausiàs March (1400–1459), who created a highly expressive though sometimes dry language laden with antitheses, metaphors, comparisons and repetitions, but very appropriate for the philosophical-moral content which he seeks to reflect; *Curial e Güelfa* (c. 1443–1450), attributable to Enyego d'Àvalos (c. 1414–1484, the grand chamberlain of Alfons the Magnanimous in Naples), which is a novel half-way between sentimental and chivalric, set in Italy and featuring an elegant tone, which shows the influence of Boccaccio's prose and of that of the early Italian humanists; *Tirant lo Blanc* (c. 1460–1464), of Joanot Martorell (c. 1410–1465), an anti-conventional and extremely realistic chivalric novel, which Cervantes called "the world's best book" and which fluctuates between Latinising cultist rhetoric and popular colloquial expression; the prose and the verses of Joan Roís de Corella (1435–1497), the former orotund and Latinising, the latter of an exquisite musicality and both with a great aesthetic beauty and an exuberant and extremely rich vocabulary; *Espill* (c. 1459–1461), by Jaume Roig (c. 1402–1478), a novel in verse of a moralistic tone and misogynous spirit, presenting a register quite close to the colloquial language of the Horta de València region; *Vita Christi* (1490), by Mother Isabel de Villena (c. 1430–1490), a vision of the life of Jesus from the standpoint of the women around him, which alternates a sumptuous Catalan full of references to the Bible and the Church Fathers, and a very rich popular vocabulary, with an abundant use of the affective resources of the language. This period is closed by the rich production of the Valencian realist poets Bernat Fenollar, Jaume Gassull and Narcís Vinyoles, among many others, who alternate religious and erotic-satirical poetry and show a great facility of versification and a remarkable mastery of linguistic resources.

In the 15th century, the most cultivated writers, such as Joan Margarit or Joan Roís de Corella, for the sake of classicism use Latin or else transfer to Catalan all sorts of Latin lexical and syntactic resources, some to the point of exacerbation. Corella presented his works as written in the 'Valencian prose style', a syntagm which we interpret today as a historiographic label designating the same Latinising style which we find in authors from Catalonia, such as Francesc Alegre, who presented his *Faula de Neptuno i Diana* as written 'in Catalan vernacular style', a similar syntagm. In the end, these are manifestations which appeared in parallel in other neighbouring languages of culture (Ferrando 1980; Nadal/Prats 1996). All things considered, the Catalan literary culture of the Late Medieval period was one of the foremost in Europe, making one of the most notable contributions.

As opposed to what has sometimes been said, the rapid decline of Catalan cultured literature at the beginning of the 16th century had nothing to do with the excesses of this rhetoricised prose but rather with the onset of the process of transculturation of an autochthonous literate nobility which took Castilian, the new court language, as a model. Indeed, from the start of the 16th century one observes a fast-paced reduction of the written output, of the cultured genres and of the works of aesthetic pretensions and linguistic quality in favour of Castilian, which was adopted by the literate nobility as a sign of social distinction and had become a prestigious international language, up to the point that it also became the court language in the independent Kingdom of Portugal (Rubió 1984–1985; Cahner 1980; Ferrando/Nicolás 2011).

7 Reflections on language

No specific grammar studies of Catalan have come down to us from the period with which we are dealing. Nevertheless, the interest in the vernacular may be observed in four aspects which imply a notable degree of linguistic reflection: a) in the treatises on troubadour precepts which in some cases, such as in the *Llibre de concordances* by Jaume March, provide us information on Catalan terms; b) in the linguistic criteria which may be deduced from the linguistic praxis of the officials of the Royal Chancery and of the leading municipal scriveneries; c) in the *grammaticae proverbiandi* and in the Latin-Catalan glossaries, which were written for the purpose of facilitating the learning of Latin, as in the case of the *Liber elegantiarum* (1472) by Joan Esteve, of the bilingual version of Antonio de Nebrija's *Lexicon* confectioned by Gabriel Busa (1507), or of the *Notes gramaticals* (1500) by Bernat Vilanova; and d) in translations. It was only towards the end of the 15th century that a rudimentary attempt was made to codify Catalan grammar with the previously mentioned *Regles d'esquivar vocables*, a text which is, furthermore, of great philological, dialectological and sociolinguistic interest.

Thanks to the formation of the European nation states, the major Western Romance languages began a process of cultivation which, with the intervention of the court humanists interested in the vernacular, led to a process of codification. Castilian, for example, already had Nebrija's *Gramática de la lengua castellana* in 1492, which was published under the auspices of Queen Isabella the Catholic. Catalan, on the other hand, lacking court support, did not undergo a similar process of codification. A series of adverse political and cultural factors (the civil war in Catalonia, the intellectual emigration to Italy, aggressions against Jews and *conversos*, the displacement of the court and of the Chancery administration to Castile under the Catholic Monarchs, etc.) prevented these humanist trends from taking root here. By contrast, the contacts of literate Catalans with Italian humanists were very fruitful. In Naples and Rome, such intellectuals as Joan Serra and Jeroni Pau were captivated by human-

ism. The *Regles d'esquivar vocables* most likely had their origin within the framework of the Roman court of the cardinal vice-chancellor Roderic de Borja (1457–1492). These were rules with which Jeroni Pau, their intellectual author, sought to endow the oral and written Catalan language with some patterns of linguistic correctness, rules which drew their inspiration from humanist parameters and which were marked by a preference for the prestigious Catalan of València (Badia i Margarit 1999; Colón/Ferrando 2011). This proposal, which was probably the fruit of the curiosity about the vernacular in the gatherings of the Roman humanists, did not come to be anything more than an erudite imagining since it could only have been feasible in a Catalan-speaking court, at that time inexistent. The “political” failure of the *Regles* was a very eloquent sign of the change of cultural paradigm which took place in the Catalan territories with the turn of the century.

8 Bibliography

- Badia i Margarit, Antoni M. (1981), *La formació de la llengua catalana*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Badia i Margarit, Antoni M. (1999), *Les “Regles de esquivar vocables” i la qüestió de la llengua*, Barcelona, Institut d'Estudis Catalans.
- Badia i Margarit, Antoni M. (2004), *Moments clau de la història de la llengua catalana*, València, Universitat de València.
- Barceló, Carmen (2011), *Àrab i català: contactes i contrastos*, València/Barcelona, Institut Interuniversitari de Filologia Valenciana/Publicacions de l'Abadia de Montserrat.
- Cahner, Max (1980), *Llengua i societat en el pas del segle XV al XVI. Contribució a la penetració del castellà als Països Catalans*, in: *Actes del Vè Col·loqui Internacional de Llengua i Literatura Catalanes*, Barcelona, Publicacions de l'Abadia de Montserrat, 183–255.
- Colón, Germà (1993), *El lèxic català dins la Romània*, València, Universitat de València.
- Colón, Germà/Ferrando, Antoni (2011), *Les “Regles d'esquivar vocables” a revisió*, València/Barcelona, Institut Interuniversitari de Filologia Valenciana/Publicacions de l'Abadia de Montserrat.
- Coromines, Joan (1971), *Lleures i converses d'un filòleg*, Barcelona, Club Editor.
- Ferrando, Antoni (1980), *Consciència idiomàtica i nacional dels valencians*, València, Institut de Filologia Valenciana.
- Ferrando, Antoni (1989), *La formació històrica del valencià*, in: *Segon Congrés Internacional de la Llengua Catalana. Àrea 7. Història de la Llengua*, València, Institut Interuniversitari de Filologia Valenciana, 399–428.
- Ferrando, Antoni (2000), *El paper dels primers editors (1473–1523) en la fixació del català modern*, *Caplletra* 27, 100–136.
- Ferrando, Antoni (2016), *La construcció de la norma cancelleresca catalana*, in: Francesc Feliu/Josep Maria Nadal (edd.), *Constructing Languages: Norms, Myths and Emotions*, Amsterdam/Philadelphia, Benjamins, 83–98.
- Ferrando, Antoni (2018), *Aportacions a l'estudi del català literari medieval*, Castelló de la Plana, Universitat Jaume I.
- Ferrando, Antoni/Nicolás, Miquel (2011), *Història de la llengua catalana*, Barcelona, UOC.
- Gabriel, Pere (ed.) (1994–1999), *Història de la cultura catalana*, 11 vol., Barcelona, Edicions 62.

- Nadal, Josep M./Prats, Modest (1982), *Història de la llengua catalana*, vol. 1: *Dels inicis al segle XV*, Barcelona, Edicions 62.
- Nadal, Josep M./Prats, Modest (1996), *Història de la llengua catalana*, vol. 2: *Segle XV*, Barcelona, Edicions 62.
- Riquer, Borja de (ed.) (1995–2000), *Història, política, societat i cultura dels Països Catalans*, 13 vol., Barcelona, Enciclopèdia Catalana.
- Rubió, Jordi (1984–1985), *Història de la literatura catalana*, 3 vol., Barcelona, Publicacions de l'Abadia de Montserrat.
- Sanchis Guarner, Manuel (1980), *Aproximació a la història de la llengua catalana*, vol. I: *Creixença i esplendor*, Barcelona, Salvat.
- Veny, Joan (1993), *Dialectologia filològica*, Barcelona, Curial/Publicacions de l'Abadia de Montserrat.
- Vernet, Joan/Parés, Raimon (edd.) (2005), *La ciència en la història dels Països Catalans*, 3 vol., Barcelona/València, Institut d'Estudis Catalans/Universitat de València.

Miquel Nicolás

12 The Origins of Modern Catalan: Cultural and Linguistic Evolution

Abstract: In the modern era, between the first quarter of the sixteenth century and the first third of the nineteenth century, the political status of Catalan underwent radical changes. Under the Habsburg monarchy, it was a declining language, culturally subordinate to Castilian. Still widely spoken in the former Crown of Aragon, where it preserved its institutions and laws and held its rank in public use, it lost part of its territory in France. Under the Bourbon rule started the political persecution of Catalan. The new dynasty designed a political, social and economic modernization program, which excluded peripheral languages from social use. Cultural production was divided into Enlightenment scholarship – expressed in Spanish, French or Latin – and Catalan popular, mostly oral culture. The Enlightenment had an ambivalent character: while showing low consideration for minority languages, it fostered archaeological scholarship and sowed the seed of Romanticism. Simultaneously, between the late eighteenth and early nineteenth centuries, the national language ideology, on which the construction of nineteenth-century nation-states would rely, started to spread. Despite all these struggles, Catalan never lost its status as a literary language – although not comparable with Spanish anymore in terms of output and prestige – and Catalan writers remained aware of the unity of the language. Hence the controversial label *Decadence* that has been applied to this period. The standard of Catalan was established essentially since then and experienced little changes in modern times beyond dialectalization.

Keywords: Habsburg monarchs, War of the Spanish Succession, Bourbons, political persecution, Enlightenment

1 The historical framework: From the transculturation to the beginning of the political persecution

In the middle of the sixteenth century, Catalan was partially excluded from the changes brought about by the advent of modernity by a sum of demographic, socio-political, economic and cultural factors. Although the very concept of “modernity” is controversial, it is undeniable that for almost three centuries, between the reign of Charles I, the first king of the House of Austria or Habsburg (1516), elected emperor of the Holy Roman Empire in 1519 as Charles V, and the beginning of the reign Isabella II

<https://doi.org/10.1515/9783110450408-019>

(1833), profound mutations took place in Western Europe, affecting the material and collective life, the orders of knowledge and the forms of exercising power.

The cornerstone that illuminated modern times was the “discovery” of the American continent. The Mediterranean ceased to be the center of gravity of European policy, which moved to the Atlantic façade. In turn, the global economy of European societies benefited from improvements in navigation and military technology, the contribution of precious metals from the New World, the opening of new trade routes and the introduction of the printing press, which revolutionized culture and introduced a new book economy.

While Christian Europe was divided and bled by the wars of religion of the sixteenth and seventeenth centuries, the cultural and linguistic weight of Islam was diluted in the Iberian Peninsula, losing its last stronghold with the eradication of the Moorish minorities of Granada and Valencia (1609). In the kingdom of Valencia, the Moriscos represented about one-third of the total population. Their expulsion was a deep social, economic and cultural crisis that the country would take decades to overcome (Barceló 1984; Casey 1979). It also led to migratory movements that ended up configuring the linguistic duality between regions of Catalan linguistic predominance and areas of Spanish-Aragonese repopulation.

In the field of ideas and aesthetic forms, Renaissance sensibility marked the struggle between dogmatic closure and submission to power, on the one hand, and the affirmation of individuality and the reform of beliefs and institutions, on the other. The doctrinal struggle between Catholic and Protestant princes was often confused with political rivalry over control of states and resources and ended up affecting the linguistic mentalities. The idea that the subject has to profess the lord's religion (*cuius regio eius religio*) was implicitly translated into linguistic communication. The tendency was thus to impose the languages with greater political support after which came, as in the case of Catalan, a recessive stage from the Renaissance onwards.

The Hispanic monarchy of the Habsburgs was a mosaic of protostates, which maintained their juridical, political, cultural and linguistic idiosyncrasy, even under the practical supremacy of the more populous and influential Castile. The Castilian matrix/mark was reinforced by several instruments of political control, among them the Inquisition, whose weight in the penetration of Castilian is an object of controversy (Ventura 1978; García Cárcel 1980). The Crown of Aragon, a confederation of sovereign territories – united by the royal dynasty, the institutions and the particular language of each territory –, was integrated into the kingdom of the Catholic Monarchs, and shortly afterwards into the vast empire of their grandson Charles I, just as Aragon, the kingdoms of Valencia and Majorca and the Principality of Catalonia preserved essentially their institutions, laws, currency, customs and traditions. We must add the use of Catalan, the language of the political, juridical and administrative documents, which inherited and modernized the koine of the medieval Royal Chancery.

This monarchy reached its peak with Philip II, the king who established Madrid as the capital of their dominions. Nevertheless, the Hispanic empire was a colossus

with feet of clay. Without a broad bourgeoisie, it was inefficient, dogmatic and unable to attend to their open military fronts. One of these, the one that it maintained with the Crown of France, was to come into an open conflict between the majority of the autochthonous institutions of Catalonia and King Philip IV with his favourite (*valido*), the count-duke Olivares (Torres 2009). This conflict, the Catalan Revolt (1640–1652), is part of the bloody Thirty Years War (1618–1648). For Catalonia, it was a failed attempt to abandon the Habsburg monarchy, similar to that of the kingdom of Portugal, which finally obtained political emancipation from Spain. The conflict ended with the annexation of ultra-/trans-Pyrenean Catalonia to France by the Treaty of the Pyrenees (1659): Roussillon and part of other ancient counties, where the primitive medieval Romance had been formed. Since then, French began to supplant Catalan in these regions, in a process of political harassment that intensified during the 20th century (Segarra 1997; Iglésias 1998).

Upon the death of Charles II, the last king of the House of Austria, a new dynastic conflict, which also led to an international struggle for European hegemony, changed the status quo between the peninsular languages. Philip V, the first king of the Hispanic branch of the Bourbons, prevailed in the War of Succession to the Spanish throne (1705–1715) over the archduke Charles of Austria. Invoking the right of conquest, the new monarch sought to modernize the structures of the monarchy, reinforcing centralism, abolishing legal and cultural singularities and imposing Castilian as the only legal instrument for government, life in society, economy and culture. The set of laws and dispositions forcefully introduced by Philip V is known as *Nueva Planta* ('Reform') and supposes the end of the previous regime of special privileges (*furs*), that had been gestating since the High Middle Ages. The *Nueva Planta* not only condemned Catalan to legal ostracism. In the south of the Valencian territory and throughout the eighteenth century, it favored the shift from Catalan to Castilian in the Baix Segura, a process that had begun with the repopulation after the eviction of the Moriscos (Montoya 1986; Mas 1994).

As successors to the Spanish throne, the Bourbons participated to the political persecution of peripheral languages. The public prosecutor of the Council of Castile, José Rodrigo Villalpando, explained in 1716 how to proceed with the imposition of Castilian. More than by force of arms, which was applied when he agreed, it was "para que se consiga el efecto, sin que se note el cuidado". That is to say, it was necessary to act with vigor of the incipient apparatuses of state, which included the institutions of knowledge, such as academies, scholars or universities (Ferrer i Gironés 1985, 22). The Enlightenment project of the Bourbons considered the renovation of a frail imperial machinery, which was approaching disintegration. The draft contained conflicting orders. During the reign of Charles III, the Council of Castile promulgated norms that prohibited the public use of languages other than Castilian. By the same time, the subjects of the former Crown of Aragon were allowed access to direct free trade with the American colonies, a measure which led to the industrial take-off of Catalonia within a few decades and, in the long run, to Catalanist claims.

The cultural break that emerged from the Renaissance deepened during the Baroque period and the Enlightenment in a series of landmarks that led the way to modernity. Latin, the true language of the European elites, persisted until the end of the eighteenth century. In this same century, the Encyclopedists conceived a vast project of systematization and criticism of knowledge. Latin, however, was doomed to disappear before the thrust of the vernaculars, which disputed the fields of politics and economics. All these contradictions crystallized in the French Revolution, whose stifled echoes reached the Iberian Peninsula during the War of Independence (1808–1814), or the French War (*Guerra del Francès*) in the Catalan historiographical tradition. This conflict laid the foundations of contemporary Spanish political identity. Far from solving the dualities between a state in construction and the survival of the *foral* period, it exacerbated them by confronting absolutism with the ideals of the liberal bourgeoisie (Vilar 1964, vol. 2, 417–423; Batllori 1997; Cahner 1998).

In short, there is a series of chained paradoxes throughout the eighteenth century and the first third of the nineteenth century. The Enlightenment, which leads directly to the revolutionary process initiated in 1789, seeks to settle accounts with the past, but sets the material and methodological bases for its study and recovery. It aims at educating the uncultivated masses, based on a supposedly rational and cosmopolitan culture, while simultaneously reaffirming the values of linguistic purism and tradition. Envisioning a more educated, freer society, it promotes the category of national languages like French or Castilian that enjoyed the prerogatives of the power; but it also lays the foundations for the Romantic rebellion and its longing for a lost past, as well as, ultimately, for the defense of minority languages like Catalan (Vilar 1964, vol. 1, 83–87; Balibar/Laporte 1974).

2 The status of modern Catalan: communicative uses and language conflicts

The historians of Catalan language and literature have often designated the long period between the middle of the sixteenth century and the beginning of the liberal era by the term *Decadència* (Comas 1986; Fuster 1976). It is true that compared to the previous and the later stage, the literary production of these three centuries is drastically reduced in quantity and quality and its dependence on Castilian cultural models is notorious. However, research has shown that this label responds to an ideological reading and does not withstand a detailed and nuanced analysis. The historical complexity goes beyond the catalogs of literary texts. And, as we have hinted, the very notion of “modernity” could be subject to critical scrutiny, which goes beyond the scope of this outline (Rossich 1989; Rossich et al. 2001; Campabadal i Bertran 2003).

Even with reservations and qualifications of a different nature, the public uses of Catalan in the political-administrative and legal documentation generally resisted the

advance of Castilian as the language that laid the foundations of the unity of the Hispanic Monarchy of the Habsburgs. The local authorities, from the viceroys to the municipal councils, through the courts and other institutions of regional law, retained Catalan as their preferred language. As for the Catholic Church, it played an ambivalent role. Apart from the liturgical uses, which remained in Latin until the Second Vatican Council, and the parish documentation, dominated by the sacramental records or *quinque libri*, the field of preaching becomes the place of a more or less fierce struggle between Catalan and Castilian (Bonet Baltà 1984; Pitarch 2000). This is particularly obvious in Catalonia, in the years preceding the secessionist revolt and during its development (1640–1652). In fact, the uprising against Philip IV, who had no support in Valencia and the Balearic Islands, gave rise to extensive political propaganda of a multilingual nature, in Spanish, Catalan and French, for and against the insurgents (Ettinghausen 1993).

This is one of the first episodes of printed political mass propaganda in Western Europe, one of the multiple sociological consequences of the cultural revolution of the printing press, whose influence had begun almost two centuries ago. In fact, the printing press favored, especially in spelling, the establishment of a normative model of prestige in every cultural and linguistic area and Catalan was no exception to this phenomenon. But in its area the culture of the book reflected the transculturation or subordination with respect to the Castilian language and culture. After all, the Spanish print market was much larger and therefore more attractive to printers in the country (Berger 1987; Fuster 1989; Chartier et al. 1996).

It is in the field of literary creation that the loss of expressive vigor of Catalan and the correlative generalization of Castilian as an instrument of author's culture became more visible (Riquer 1964, vol. 3; Comas 1964, vol. 4; Nadal 1983). The literary products written in Catalan followed the aesthetic models, the themes and formal motives of the Castilian Baroque. They were conceived mainly for the consumption of the elite, but also, particularly through theatrical performances, increasingly for urban populations, whose literacy rates were still very low. Poetry of various genres stands out from devotional to satirical production. As far as prose is concerned, the texts produced during the three centuries that occupy us here are by no way comparable, in aesthetic tension, to the classics of the fifteenth century. However, there are some works in prose that embody the model of the Renaissance dialogue. The best example is a book that remained unpublished until the nineteenth century, *Los col·loquis de la insigne ciutat de Tortosa* (1557), by Critòfor Despuig (Duran/Solervicens 1996). The apologetic literature must be added to this genre, particularly in the subgenres of "the defenses of the language" and historiographical exposition (Pitarch 1972; Baró i Queralt 2009; Feliu et al. 1992).

However, in spite of political pressure, demolinguiistic inferiority and cultural dependence, Catalan survived in oral use and in an extensive list of paraliterary written genres, from memorialism to traditional romances in verse (in Spanish *literatura de caña y cordel*), dialogues, collections of proverbs, and various forms of

popular religiosity, such as *goigs* ('songs in honor of Mary or a saint'), devout books and catechisms, the lives of saints or the abundant displays of Marian devotion (Campabadal i Bertran 2003).

3 The destandardization: weakening of koine, convergence with Spanish, increasing dialectalization

In the sixteenth century Catalan had left behind the processes of phonetic and morphosyntactic change that led from the Vulgar Latin spoken in the Tarraconensis to an evolved Romance system. Political vicissitudes had kept it out of the Occitan space, although it continued to maintain close contact with the latter, and in the literary texts of the period, certain reminiscences of the troubadour lyric are found. It is difficult to reconstruct the physiognomy of oral language from writing, especially since the *scripta* of Catalan – which remained very uniform until the end of the fifteenth century – dispersed from the sixteenth century and experienced a quantitative and qualitative decline. In any case, the oral and written manifestations of Catalan between the sixteenth and the eighteenth centuries reflect the phenomena of linguistic interference and diatopic variation coinciding with the decline of socio-political status of language.

Cultural subordination to Castile accentuated the process of lexical hispanicization. As seen in 711 The Growth and Expansion of Catalan (1213–1516), this consisted in the fact that Gallo-Romance lexemes were replaced by their Ibero-Romance counterparts (Colón 1993).

During the modern period, the Castilian sociopolitical and cultural pressure went further and resulted in the massive entry of castilianisms into all lexical fields: *agasajo* 'reception', 'entertaining', *alabar* 'to praise', *apoiar* 'to support', *aposeno* 'accommodation', *aplausos* 'applause', *assentarse* 'to settle down', 'to sit down', *carinyo* 'affection', *concejals* '(town) councillors', *cotxe* 'carriage', *cuidar* 'to look after', 'to care for', *entresuelo* 'raised ground floor', *llàstima* 'pity', *loco* 'mad', *membrillo* 'quince', *resar* 'to pray', *tonto* 'stupid', 'silly', *trago* 'sip', etc. The period of greatest interference occurred between the *Nueva Planta* (1715) and the first attempts at language purification in the middle of the nineteenth century (Bruguera 1985, 67).

On the other hand, the three great regions of Catalan, the Principality of Catalonia and the kingdoms of Valencia and Majorca, entered a long period of disintegration and loss of political representativeness (Sanchis Guarner 2002; Massot 1985; Martínez 2000), which saw the dissolution of institutions that provided shared linguistic models for written usage. The persistence of the official status of Catalan after the *Decretos de Nueva Planta* (1707–1716) contributed to the maintenance of the cultured use of the language in the administrative field. The memory of this administrative

language (acts of Parliament, ceremonial books, etc.) contributed decisively to the partial recovery of the cultured *scripta* through the nineteenth century by grammarians and lexicographers.

Without a common pattern in the *scripta* and under the pressure of centripetal forces, each territory tended to reinforce its regional peculiarities. In Eastern Catalan, the tendency to neutralize the atonic vowels [o], [ɔ] and [u] (*rodó* [ru'do] 'round') and of [a], [e] and [ɛ] in the neutral vowel [ə] (*pare* ['parə] 'father') was accentuated. Dialectalization also affected the written texts, in which formal variants (*genoll*, general Catalan, *ginoll*, Valencian), geosynonyms (*noi*, Catalan/*al·lot*, Balearic/*xic*, Valencian), morphosyntactic structures (*la hi donava* 'he/she gave it (fem.) to him/to her/to them', general Catalan/*li la donava*, Valencian) and phraseological patterns (*Això rai* 'It's not so bad, no matter/problem' *Déu n'hi do* 'and how!', used only in Catalonia) became regional innovations mutually exclusive in each dialect.

4 Linguistic reflection: Survivals of idiomatic awareness and shrinkage of grammar

The introduction of printing in the Catalan-speaking domain (1473) raised the need to standardize the spelling, since printed books catered to a much broader readership than manuscript books did. The standardization of writing, however, was not limited to spelling but involved a reconsideration of the formal levels and possibilities of written expression. It reopened the controversy over the model of prestige language, the only one that was considered apt for literary communication. Rhetoric teachers framed this debate in terms of the Latin contrast between *sermo rusticus* and *sermo urbanus*. In addition, the debate, put forward everywhere with relatively similar arguments and emphases, echoed what has been called subsequently and with reference to Dante *la questione della lingua*, that is to say, the claim to use the language of the people, which was comparable in aptitude and dignity to Greek and Latin.

The linguistic reflection on Catalan between the sixteenth century and the first third of the nineteenth century is sparse, as corresponds to the decline of Catalan expression in the public space. During the late Renaissance and the Baroque period, it is basically circumscribed to spelling lists of words and lexicographical works that actually pursue the teaching of Latin, Spanish or French. The first worth dictionary is due to Pere Labèrnia, a Valencian based in Barcelona who inspired all the lexicographers of the nineteenth century.

The grammarians of the Enlightenment (Pere Màrtir Anglès, Josep Ullastrà, Carles Ros, Joan Petit Aguilar, etc.) provided, with unequal preparation, a description of the cultured use of the variety they knew best. The most interesting grammatical contributions are those of Josep Pau Ballot and the Minorcan Antoni Febrer i Cardona. It is

worth noting that – compared to the rest of the Catalan-speaking territories – Minorca enjoyed a situation of limited political pressure, at least during the three periods during which the British occupied the island (Segarra 1985; Ginebra 1996). It is not a mere coincidence that Joan Ramis i Ramis, the principal Catalan-writing playwright of the period, was also a Minorcan.

At the beginning of the eighteenth century, Spanish was endowed with a normative institution, the Real Academia Española (1713). Under the impulse of Enlightenment, Spanish became with Latin, and to a lesser extent with French, a language of knowledge and scientific communication. Due to their stigmatization, the transmission of literary memory in other peninsular languages came to an end. According to some writers, these languages were characterized by a subaltern condition or a congenital incapacity for ideas. In fact, the Enlightenment culture examined suitable topics for an elaborate communication on language while preparing the way for modern science. And this lack of security spread even among Catalan authors. Thus, Antoni de Capmany, one of the scholars whose outstanding achievement was his research into the Catalan history and literature of medieval times, considered however that Catalan was an “idioma provincial muerto hoy para la república de las letras” and thus inadequate for modern cultural expression. Some authors, who took refuge in their native culture, reacted against this elitist vision and reclaimed their own tradition, announcing the preromantic sensibility.

The Enlightenment attitudes, and the linguistic practices they inspired, appeared for the first time in the *Real Cédula* of Aranjuez (1768) which ordered the teaching of Castilian as “el idioma general de la Nación” (Ferrer i Gironès 1985, 37; cf. Lüdtke 1989, 269–271), and culminated in the Jacobin ideology of the French Revolution, which elaborated the notion of “national language” (Balibar/Laporte 1974). During the occupation of Catalonia, the French tried to gain the adhesion of the Catalans by granting their language the rank of co-official language in some contexts, but the Catalans were no longer trained to write documents in their own language and used Spanish instead. They simply did not recognize French as their universal language, and considered their language none of the French’s business, since their national identity was Catalan and Spanish (Kailuweit 1997, 196–206). The truth is, however, that the Napoleonic armies began to propagate the Jacobin idea that modernity and progress were associated with the language of the new national state. This ideology, exported to half Europe, proved, for various reasons, more effective in the imposition of a homogeneous language than the Enlightenment project. The relative failure of this project was due to the fact that it barely caught on a minority of scholars, without sufficient resources for literacy and mass control, something that will not come until the industrial revolution.

In modern times, the dialectalization of Catalan affected the awareness of an ascription to the same linguistic trunk. As the perceived unity of Catalan dialects declined, onomastic particularism was reinforced when it came to naming the common language (Schmid 1988). In fact, particularistic glottonyms (*llengua catalana*,

llengua valenciana, *llengua mallorquina*) was clearly the result of a desire for political self-assertion that was very strong among the ruling classes of the young kingdom of Valencia. The political context, however, was very different from the one that emerged in the sixteenth century. Now, a literate, humanistically educated minority was propagating a new denomination for the Catalan of literary texts, *llemosí* or *llengua llemosina*. The metalinguistic use of this name is documented in many texts, parallel to the other denominations that were still being used. Actually *Limousin* is a diatopic variety of northern Occitan. By applying the term to the common language, they tried to avoid anti-Catalan suspicions, present already in the initial times of the Hispanic Monarchy. At the same time, they tried to preserve, but in an unclear form, the awareness of the unity of the language. The conceptual confusion lasted during the entire modern era and became strongly rooted among the Valencian writers who, even at the beginning of the twentieth century, when political Catalanism began, continued to cling to *llemosinisme* (Rafanell 1991). A proof of dialectalization can be seen in the proliferation of interdialectal prejudices, especially among Valencian scholars, who tend to present their speech as sweeter, softer or more elaborate than the Catalan of Catalonia, which is considered rougher, ruder and more neglected.

5 Bibliography

- Badia i Margarit, Antoni M. (1999), *Les "Regles d'esquivar vocables" i "la qüestió de la llengua"*, Barcelona, Institut d'Estudis Catalans.
- Balibar, René/Laporte, Dominique (1974), *Le français national: politique et pratique de la langue nationale sous la Révolution*, Paris, Hachette.
- Balsalobre, Pere/Gratacós, Joan (edd.) (1995), *La llengua catalana al segle XVIII*, Barcelona, Quaderns Crema.
- Barceló, Carmen (1984), *Minorías islámicas en el País Valenciano. Historia y dialecto*, Valencia, Universidad de Valencia.
- Baró i Queralt, Xavier (2009), *La historiografia catalana en el segle del Barroc (1585–1709)*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Batlloori, Miquel (1997), *Obra completa*, vol. 9: *La Il·lustració*, València, Tres i Quatre.
- Berger, Philippe (1987), *Libro y lectura en la Valencia del Renacimiento*, València, Institució Alfons el Magnànim.
- Bonet Baltà, Joan (1984), *L'Església catalana, de la Il·lustració a la Renaixença*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Bruguera, Jordi (1985), *Història del lèxic català*, Barcelona, Enciclopèdia Catalana.
- Cahner, Max (1998), *Literatura de la Revolució i la Contrarevolució (1789–1849)*, vol. 1, Barcelona, Curial.
- Campabadal i Bertran, Mireia (2003), *El pensament i l'activitat literària del Setcents català*, 2 vol., Barcelona, Universitat de Barcelona.
- Casey, James (1979), *El Regne de València al segle XVII*, Barcelona, Curial.
- Chartier, Roger, et al. (1996), *La cultura escrita a la Catalunya Moderna*, L'Avenç 199, 25–47.
- Colón, Germà (1993), *El lèxic català dins la Romània*, València, Universitat de València.
- Comas, Antoni (1964), *Història de la literatura catalana*, vol. 4, Esplugues de Llobregat, Ariel.

- Comas, Antoni (1986), *La Decadència*, Barcelona, La Llar del llibre.
- Duran, Eulàlia/Solervicens, Josep (edd.) (1996), *Renaixement a la carta*, Barcelona/Vic, Universitat de Barcelona/Eumo.
- Ettinghausen, Henry (1993), *La guerra dels segadors a través de la premsa de l'època*, 4 vol., Barcelona, Curial.
- Feliu, Francesc, et al. (1992), *Tractar de nostra llengua catalana. Apologies setcentistes de l'idioma del Principat*, Vic, Eumo/Universitat de Girona.
- Ferrando, Antoni/Nicolás, Miquel (2011), *Història de la llengua catalana*, Barcelona, UOC.
- Ferrer i Gironés, Francesc (1985), *La persecució política de la llengua catalana*, Barcelona, Edicions 62.
- Fuster, Joan (1976), *La Decadència al País Valencià*, Barcelona, Curial.
- Fuster, Joan (1989), *Llibres i problemes del Renaixement*, València/Barcelona, Institut de Filologia Valenciana/Publicacions de l'Abadia de Montserrat.
- García Cárcel, Ricardo (1980), *Herejía y sociedad en el siglo XVI. La Inquisición en Valencia, 1530–1609*, Barcelona, Península.
- Ginebra, Jordi (1996), *L'obra gramatical d'Antoni Febrer i Cardona (1761–1841)*, Maó, Institut Menorquí d'Estudis.
- Iglésias, Narcís (1998), *La llengua del Rosselló, qüestió d'estat*, Vic/Girona, Eumo/Universitat de Girona/Universitat de Vic.
- Kailuweit, Rolf (1997), *Vom EIGENEN SPRECHEN. Eine Geschichte der spanisch-katalanischen Diglossie in Katalonien (1759–1859)*, Frankfurt am Main, Lang.
- Lüdtke, Jens (1989), *Acerca del carácter imperial de la política lingüística de Carlos III*, in: Günter Holtus/Georges Lüdi/Michael Metzeltin (edd.), *La Corona d'Aragó i les llengües romàniques. Miscel·lània d'homenatge per a Germà Colon*, Tübingen, Narr, 267–274.
- Martínez, Catalina (2000), *La llengua catalana a Mallorca. Segle XVIII i primer terç del XIX*, Barcelona, Publicacions de l'Abadia de Montserrat/Universitat de les Illes Balears.
- Mas, Antoni (1994), *La substitució lingüística del català (l'administració eclesiàstica d'Elx a l'edat moderna)*, Alacant, Diputació Provincial.
- Massot, Josep (1985), *Els mallorquins i la llengua autòctona*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Montoya, Brauli (1986), *Variació i desplaçament de llengües a Elda i a Oriola durant l'Edat Moderna*, Alacant, Institut d'Estudis "Juan Gil-Albert".
- Nadal, Josep Maria (1983), *"Usar de llenguatge artificios" en el segle XVI: ideologia lingüística i llengua literària*, in: *Actes VIè Col·loqui Internacional de l'AILLC*, Barcelona, Publicacions de l'Abadia de Montserrat, 89–126.
- Pérez Saldanya, Manuel (2013), *Del llatí al català. Morfosintaxi verbal històrica*, València, Universitat de València.
- Pitarch, Vicent (1972), *Defensa de l'idioma*, València, Tres i Quatre.
- Pitarch, Vicent (2000), *Llengua i predicació durant el Barroc valencià*, Barcelona/València, Publicacions de l'Abadia de Montserrat/Institut de Filologia Valenciana.
- Rafanell, August (ed.) (1991), *Un nom per a la llengua. El concepte de llemosí en la història del català*, Vic, Eumo.
- Rafanell, August (1999), *La llengua silenciada. Una història del català, del Cinc-cents al Vuit-cents*, Barcelona, Empúries.
- Riquer, Martí de (1964), *Història de la literatura catalana*, vol. 3, Esplugues de Llobregat, Ariel.
- Rossich, Albert (1989), *Renaixement, manierisme i Barroc en la literatura catalana*, in: *Actes del Vuitè Col·loqui Internacional de llengua i literatura catalanes*, vol. 2, Barcelona, Publicacions de l'Abadia de Montserrat, 149–179.
- Rossich, Albert, et al. (edd.) (2001), *El teatre català dels orígens al segle XVIII*, Kassel, Reichenberger/Universitat de Girona.

- Sanchis Guarner, Manuel (2002), *Els valencians i la llengua autòctona durant els segles XVI, XVII i XVIII*, València, Universitat de València.
- Schmid, Beatrice (1988), *Les traduccions "valencianes" del "Blanquerna" de Llull (València, 1521) i de l'"Scala Dei" d'Eiximenis (Barcelona, 1523)*, Barcelona, Curial/Publicacions de l'Abadia de Montserrat.
- Segarra, Mila (1985), *Història de l'ortografia catalana*, Barcelona, Empúries.
- Segarra, Mila (1997), *El conflicte lingüístic castellà-català als segles XVI i XVII*, in: Pere Gabriel (ed.), *Història de la cultura catalana*, vol. 2: *Renaixement i Barroc*, Barcelona, Edicions 62, 167–192.
- Torres, Xavier (2009), *Naciones sin nacionalismo. Cataluña en la monarquía hispánica (siglos XVI–XVII)*, València, Universitat de València.
- Ventura, Jordi (1978), *Inquisició espanyola i cultura renaixentista al País Valencià*, València, Tres i Quatre.
- Vilar, Pierre (1964–1968), *Catalunya dins l'Espanya moderna*, 4 vol., Barcelona, Edicions 62.

Jenny Brumme

13 *Renaixença*

Abstract: This chapter provides an overview of the *Renaixença* movement, that is, the revival of Catalan language and literature in the second half of the 19th century. For a better understanding of the linguistic, cultural and socio-political changes associated with the movement, this article also describes the conditions in the early 19th century in terms of break and continuity. This includes the “roofing” of a mainly monolingual Catalan society by Spanish, the prestigious but foreign language, more closely linked with formal and written usage. The construction of the modern Spanish state and its ongoing centralisation are documented through the series of laws restricting the use of Catalan. This chapter’s main focus is the recovery of the language for literary use through the *Jocs Florals* poetry competition which was reinstated in 1859. The second focus regards the use of Catalan in the press. The late 19th century also saw the emergence of political Catalanism, a key ideological component of which became the nurturing of a separate language. As references to a common writing tradition had been lost over the previous centuries, the reintroduction of Catalan for most written uses triggered a debate over which model should be adopted for standardisation: one concerned with the past or one that took as its reference point the use of spoken Catalan at the time.

Keywords: language revival, literary use, print-media, language awareness, code conflict

1 Origin and meaning of the term *Renaixença*

The concept of *Renaixença*, that is to say, the revival of Catalan language and literature in the 19th century, is today a controversial notion (cf. Marfany 2001, 465–492). The term should not be confused with the *Renaissance*, the European Renaissance of classical ideas after the Middle Ages. The term *Renaixença* reflects the anthropomorphic view of language development starting with the birth of a language (or literature), followed by its golden age, its decline (formerly known as *Decadència* in Catalan) and either its revival or death (cf. Domingo 2009, 216).

Recent research has questioned this anthropomorphic model, arguing that the use of Catalan as a spoken language never really declined, therefore making the term *Renaixença* a misnomer. What is clear, however, is that the mid-1800s did see a surge of interest in the renewal of Catalan literature and the recovering of Catalan as a language for the literary (written and formal) use. Contemporary witnesses in the 1860s and 1870s expressed this idea through terms such as: *despertament* (‘awaken-

<https://doi.org/10.1515/9783110450408-020>

ing'), *desvetllament* ('arousal'), *renaixement* or *renaixensa* ('rebirth'), *resurrecció* ('resurrection'), *restauració* ('restoration'), *revifalla* ('revival') or *reviscolança* ('resurgence') (Duran i Tort 2001, 51–52). While these terms commonly appeared in collocates such as literary, artistic, industrial or scientific revival (Duran i Tort 2001, 52), what the Catalan language needed, however, was to be awoken, not revived.

The consolidation of the term *Renaixença* during the 1890s, written in uppercase and without qualifying adjectives, may have been influenced by the prestigious publishing company "La Renaixensa" (Duran i Tort 2001, 53), which published a fortnightly magazine of the same name (1871–1880), which was later converted into a political newspaper (1881–1905). In its first period, the magazine provided a platform for those parts of Catalan society that shared a love for their homeland and for the Catalan language (Duran i Tort 2001, 207).

The *Renaixença* movement shows similarities with other attempts to renew language and literature, such as the Italian *Risorgimento*, the Galician *Rexurdimento* or the Occitan *Felibrige*, with which the Catalan movement shared a rather unfruitful and conflict-laden relationship.

In this first approximation of the concept of *Renaixença* it should be mentioned that, in parallel with the Romantic idea of language, the idea of *patria* was also taking shape. Therefore, it is no surprise that different political movements deeply linked with the different linguistic manifestations appeared and formulated their objectives about the place of Catalonia and other Catalan-speaking areas within the framework of the different nation states to which they belonged.

2 The language policy of the modern state

The revival of Catalan language and literary culture coincides with the construction of the modern state in Spain, as well as in France and Italy where Catalan was also spoken. During the Spanish War of Independence (1808–1814) the 1812 Constitution represented the first liberal programme in Spain and established a limited hereditary monarchy with a government of elected representatives and based on the division of power. Although the liberal programme included some fundamental achievements such as freedom of thought and of the press, its centralism and demands for national unity clearly reflected the programme's Jacobin legacy. Until the Spanish Constitution of 9th December 1931, none of the earlier constitutions or constitutional projects echoed the linguistic diversity of Spain.

From the beginning of the 19th century until the First Spanish Republic (1873–1874), there were several revolutionary periods followed by periods in which the previously implemented reforms were recalled and the old regime is reinstalled. Despite these political upheavals, towards the end of the reign of Isabel II (1833–1868) liberalism was properly implemented and the centralist and unitary character of the state improved. This period of political instability, but economic expansion gave way to the

Bourbon Restoration (1875–1931), which consolidated the two-party system and fostered the values of bourgeois culture. At the end of the 19th century the former mainly monolingual society shifted definitively into a society where use of the language fulfilled the criteria of diglossia. This means that the use of Spanish (or French, respectively) was associated with formal communication (in the fields of political and legal, scientific-technical as well as socioeconomic communication), while Catalan was connected to the symbolic space of every-day expression and intimacy, as well as certain genres of high and popular literature (Ferrando Francés/Nicolás Amorós 2016, 304).

Throughout the entire 19th century several measures were taken by the central government which solidified the unification of Spain and strengthened the diglossic distribution of language usage. After the enactment of the *Nueva Planta* Decrees by which Philip V (1700–1724) suppressed the institutions, privileges and ancient charters in all parts of the former Crown of Aragon (→12 The Origins of Modern Catalan: Cultural and Linguistic Evolution), other measures were established to support the construction of a centralised state and a homogeneous internal market. Among the multiple changes introduced in this period the most significant include the creation in 1844 of the *Guardia Civil*, a military force with the purpose of restoring order and protecting property, and the division of Spain into 49 provinces (1833) following the model of the French departments. After joining the Latin Monetary Union on the eve of the 1868 Revolution, the currency unit of the peseta and the decimal system of measurement were introduced.

In the second half of the 19th century, centralistic legislation was directly imposed on language use. The most important law due to its long-term impact was the *Ley de Instrucción Pública* or *Ley Moyano* (1857). This law made elementary education compulsory and free for all children attending public schools (article 7). Subjects to be taught in schools included, among others, Spanish language with spelling exercises (article 2), for which the only textbooks permitted were the grammar and orthography published by *Real Academia Española* (article 88). By this means, Spanish was set as the sole language of education.

There followed other laws concerning the civil code, which provoked protest from Catalan civil society. Firstly, the *Ley del Notariado* (1862) decreed that all public documents had to be written in Spanish (article 25). Secondly, the *Ley del Registro Civil* (1870) required all municipal courts to create a civil registry (article 2) thereby invalidating parish records and language used in them. The use of Spanish for marriage ceremonies was explicitly mandated (article 58). Finally, the *Ley de Enjuiciamiento Civil* (1881) required all documents be produced in Spanish (article 601).

Returning to the education system, the *Decreto Romanones* (1902), so named after the man who introduced it, Álvaro Figueroa, the Count of Romanones (1863–1950), is also worth mentioning. The decree made it a punishable offence to teach the catechism in languages other than Spanish (article 2). The “national language” – so the argument went – should be known by all Spanish citizens and a lack of this knowledge would mean serious harm to the national interests (“los altos intereses de la Patria”).

3 Catalan in the early 19th century: break and continuity

Unlike the use of written Catalan in formal registers and genres, there are more than a few eyewitness testimonies for the healthy state of spoken Catalan at the beginning of the 19th century. Foreigners who visited Catalonia bear witness to the fact that Catalans largely spoke their own language, using Spanish either to a lesser extent or not at all. One such visitor was Alexandre Laborde (1773–1842) who travelled across the peninsula between 1800 and 1805 working as an archaeologist. He said of the Catalan language:

“Les Catalans ont une langue qui leur est particulière [...]. On la parle dans toute la Catalogne avec beaucoup de variations, suivant les divers cantons, avec plus de pureté sur les montagnes, avec plus d’altérations dans les grandes villes. La préférence nationale du Catalan lui fait préférer sa langue à celle des Espagnols; aussi le castillan est-il fort peu en usage dans la Catalogne; et, lorsqu’on l’y rencontre, il est défiguré et méconnaissable par le mélange d’expressions et de tournures catalanes” (Laborde 1827, 191–192).

While this quotation focuses more on Catalan and the language skills of the Catalans, the following quotation makes a more political assessment. Another famous traveller Richard Ford (1796–1858) visited Catalonia in 1831 and in 1845 published his book in which he attests to the otherness of Catalans and Catalonia: “The Catalans are neither French nor Spaniards, but a distinct people, both in language, costume, and habits [...]. Catalonia, [...], is the strength and weakness of Spain; and no province of the unamalgamating bundle which forms the conventional monarchy *de las Españas* hangs more loosely to the crown than this classical country of revolt, which is ever ready to fly off” (Ford ³1855, 392–393). According to these and many other eyewitness accounts, despite the restrictive legislation of the Spanish Crown (1809 Language Law and Language Policies), the language that most of the population acquired spontaneously at home was Catalan. Nevertheless, the influence of Spanish only continued to grow, not because it was taught at school, but rather due to its social and literary prestige. The status of Spanish as the language that provided access to knowledge and intellectual exchange should not be underestimated. This might have inspired the frequently quoted verdict of Antonio de Capmany y de Montpalau (1742–1813), who saw Catalan as an “idioma antiguo provincial, muerto hoy para la República de las letras, y desconocido del resto de Europa” (1779, vol. 2, *Apendice de algunas notas*, 54). Therefore, in his in-depth reconstruction of the discourse on language in Catalonia in the early 19th century Kailuweit (1997) distinguishes between four types of ‘learned languages’ (*grammolects*) in contrast to the language transmitted at home (*genolect*): standard Spanish, regional Spanish, Spanish-influenced Catalan and traditional written Catalan (Kailuweit 1997, 124; 1999).

It appears that Catalan occupied a low status in this regard, while standard Spanish enjoyed a high status (diglossia). The lack of standardisation and the ongoing

fragmentation of traditional spelling into dialectally marked practices contributed to this devaluation. One of the tasks of the *Renaixença* movement in the second half of the century was to readjust the functional distribution between Catalan and Spanish, not in terms of a break, but a ‘sudden linking’ of the concepts of language and the development of a social identity (Kailuweit 1997, 114). However, the process of *Überdachung* (Kloss ²1978) or ‘roofing’ of Catalan by Spanish continued throughout the whole century (cf. Brumme 1997, 58–80).

At the end of the 18th century, but mainly during the Spanish War of Independence (1808–1814), the politically motivated use of Catalan initiated a break in the perception of the language and in the longer term led to essential changes in language awareness. One of the first signs of this awareness can be found in the *Gramatica i apologia de la llengua catalana* (1814), compiled in the middle of the Napoleonic invasion by Josep Pau Ballot (1747–1821). In his bilingual (Spanish and Catalan) dedication to the Spanish Royal Board of Commerce of Barcelona, Ballot, a priest, professor of rhetoric and author of several didactic works on Spanish and Latin, claimed that his grammar “pretende axáltar la lengua cathalana y elevarla al mas alto grado de perfeccion” (1987, VI). On several occasions in the book, Ballot stresses that Catalan is a language in its full sense and not a dialect or jargon (Ballot 1987, XV and XXIX). He also viewed Catalan as a different language from Provençal or Limousin (Ballot 1987, 259). His grammar tried to systematise the different orthographical and morphological solutions that existed in parallel at the time. In his attempts to connect the contemporary usage of Catalan to the writing traditions of the 16th and 17th century, Ballot is to some extent in line with the advocates of ‘academic Catalan’ (Segarra in Ballot 1987, [29]; cf. 6).

Nevertheless, Ballot also poses the relevant question of whether it is worthwhile ‘cultivating the Catalan language if the language of the entire nation is Spanish’ (Ballot 1987, XXIV). Throughout the 19th century, the authors of grammars and dictionaries both for Catalan and Spanish edited in the Catalan-speaking areas would return to this question and more than a few authors endeavoured to provide resources for learning Spanish (Schmid 2014). There are distinct reasons for this, but one was to facilitate learning Spanish from the basis of Catalan as a mother tongue, which ultimately attests to the uninterrupted presence of the language in society. This can be seen in the *Gramática catalana-castellana* (1847) by Magí Pers y Ramona, which set out to introduce Spanish through Catalan. According to Pers y Ramona, Catalans urgently needed to learn the ‘general language of the Spanish nation’ (Pers y Ramona 1847, 6) and with a little study they would be able to do so. This trend of promoting the acquisition of Spanish was countered by the trend of highlighting the drawbacks of writing in Catalan, with several authors criticising the lack of standardisation (cf. 6).

Another source of difficulty is discussed in the foreword to the first novel in Catalan. Antoni de Bofarull (1821–1892) (cf. 4.1) pointed directly to the lack of writing habits and the absence of a regulating authority to standardise the orthographical chaos. Over the years he would repeat and elaborate on this type of remark: [Since

Catalan] ‘stopped being official, it has been written in many ways, with each author or publisher adopting the system that he thinks or knows is the best’ (Bofarull 1862, 10). In this context, it does not come as a surprise that the restitution of writing practices in Catalan was accompanied by several controversies surrounding spelling and morphology through the second half of the 19th century.

4 The recovery of written usages

4.1 Usage in poetry and literature

Traditionally in the case of Catalan, the decline of the language (*Decadència*) has been identified as the period from 1500 to the early 1800s, while the *Renaixença* was seen to start in 1833 with the publication of the poem *La pàtria. Trobes* (‘The Homeland. Minnesongs’) by Bonaventura Carles Aribau (1798–1862). This poem, published in the newspaper *El Vapor* was later renamed *La pàtria* (‘The Homeland’) and is generally known as *Oda a la Pàtria* (‘Ode to the Homeland’).

However, the impact of this poem has always been overestimated, as has Aribau’s contribution to the *Renaixença*. Indeed, this appraisal of the ode was target of criticism in one of the first extensive outlines of the *Renaixença* movement. In his *Historia del Renacimiento literario contemporáneo en Cataluña, Baleares y Valencia* (1880), Francisco María Tubino (1833–1888) argued that, though the ode is seen as the starting point of the movement, there are not sufficient grounds for this assessment (Tubino 2005, 1880, 162; Domingo 2005, 41). This is in line with the opinion of more recent research that stresses Aribau’s lack of commitment to Catalan in his literary work, which is mainly written in Spanish (Rafanell 1999, 131). In fact, it is worth mentioning the point made by Hina that the ode can neither be seen as a parody of Romantic lyrics nor as a declared attachment to Catalan as a mother tongue (Hina 1978, 107; 1986, 119). Nevertheless, the ambivalence of this lyrical masterpiece has contributed to perpetuating the traditional view of this poem as the starting point of the *Renaixença*.

The real beginning of the *Renaixença* should be situated somewhere between 1836 and 1839, when Joaquim Rubió i Ors (1818–1899) started publishing his first poems. In 1839, the first of what would later become his most popular poems came out under the pseudonym *Lo Gayter del Llobregat* (‘The Piper of the Llobregat’). From February 1839 to November 1840 the *Diario de Barcelona* published 19 of these poems, which in 1841 were collected in the volume of the same name. The foreword to this volume would prove highly significant for his contemporaries. As well as advancing both the metrical and linguistic basis of Catalan poetry in the latter half of the 19th century, these poems are also set against an unmistakably Catalan backdrop. Instead of the troubadours and their world, these poems are concerned with the piper, Barcelona, the *Llobregat* river, the local craft world, and popular culture in the guise of high literature.

In 1863 Rubió i Ors was awarded the title of *Mestre en Gay Saber* ('Master in the Art of Poetry') and in 1890 was nominated president of the *Jocs Florals*. Despite his commitment to Catalan poetry and the literary revival he published and would continue to publish his works as an educationalist and biographer in Spanish. This is in line with the contemporary use of Catalan, which probably reached the low point in its recovery in the 1830s and 1840s before gradually rebuilding genre by genre up until the end of the century and beyond.

The *Jocs Florals* poetry competition was reinstated in Barcelona in 1859. The title harks back to medieval 'floral games' (→14.1 The Language Reform, the Institut d'Estudis Catalans and the Work of Pompeu Fabra). The *Renaixença* movement was nostalgic for the glorious heyday of the Catalan empire in the Middle Ages, when Catalan political power and the Catalan language were at their zenith. The initiative in 1859 was followed by others in the other Catalan-speaking areas and contributed to restoring the usage of Catalan in literary (formal and written) contexts.

It has always been stressed that the *Renaixença* movement emerged under the influence of Western European Romanticism, particularly such medieval-style authors as Walter Scott (1771–1832) and Victor Hugo (1802–1885). Although the primary form of literary expression in Catalan was in verse, Catalan literature shared many of the main motifs and topics of European Romanticism (Ferrando Francés/Nicolás Amorós 2016, 316). The ongoing expansion of the literary use of the language gave rise to the full development of poetry, theatre and narrative fiction in the 1870s and 1880s thanks to authors such as Jacint Verdaguer (1845–1902), Àngel Guimerà (1845–1924) and Narcís Oller (1846–1930). The most difficult literary form to recover proved to be that of the novel, because the Catalan writing tradition of this type of prose had been lost in the Middle Ages and the new literary genres and trends of the 19th century were introduced via Spanish translations (Tayadella 2012, 98–99). The genre of the novel was not re-established in Catalan until the publication of *La Orfaneta de Menargues* ('The Orphan Girl from Menargues', 1862) by Antoni de Bofarull. As Bofarull stated in the prologue to the novel, an impoverished vocabulary and lack of standardisation greatly hampered this revival of the genre in Catalan (Tayadella 2012, 104). Thanks to the predominant literary trends of Naturalism and Realism writers such as Oller consolidated the credibility of narrated facts and speech by describing and evoking everyday reality in Catalan.

The council of the *Jocs Florals* conceded the title of *Mestre en Gay Saber* to Guimerà in 1877 and to Verdaguer in 1880. There were three categories in the contest: *Patria*, *Fides*, *Amor* ('Country, Faith, Love'). The title *Mestre en Gay Saber* was awarded if a poet won all three prizes for a patriotic poem, a love poem and a religious poem. Verdaguer's contribution to the revival of Catalan literature cannot be overstated. Among his extensive output figure two important epics. At the time, the general idea prevailed that a literature had to possess an epic, that is to say, an artistic and linguistic referent, in order to count among the other national literatures (Pinyol i Torrents 2009, 219). The first epic *L'Atlàntida. Poema* (1878) formed the basis for

Verdaguer's international reputation as it chants the mythical origins of the Iberian Peninsula and not those of a specific nation. The second epic *Canigó* (1886), which takes its name from the Pyrenean peak, articulates in a singular and metrically diverse way the birth of Christian Catalonia. In its entirety, this work constitutes the culmination of Catalan epic poetry in the 19th century.

The *Renaixença* reached its zenith around 1890 (Ferrando Francés/Nicolás Amorós 2016, 329), when the new artistic, literary and architectonic movement of Modernism took root. It was also during the late 19th and early 20th century that the first attempts to codify Catalan were made. A group of Catalan intellectuals from the modernist journal *L'Avenç* (1881–1893) started to think about how to bring an end to the prevailing (ortho-)graphic anarchy through a new codification of the language (Ferrando Francés/Nicolás Amorós 2016, 492–495). The beginning of the 20th century was marked by two events of major importance: the *Primer Congrés Internacional de Llengua Catalana* ('First International Congress on the Catalan Language', 1906) and the foundation of the Institut d'Estudis Catalans (1907) as an "academic, scientific and cultural corporation the aim of which is scientific research, principally that involving all the elements of Catalan culture" (IEC 1992, 5).

4.2 The use of Catalan in the press

One of the most significant indices for the recovery of Catalan in the written domain is its use in the press. In contrast to the elevated and sophisticated, but often hackneyed use of language in the poetical creations of the *Jocs Florals*, the press was able to reflect the more everyday use of the language. Although Spanish-speaking periodicals prevailed throughout the entire 19th century, following the restoration of the *Jocs Florals* the inclusion of Catalan compositions became more frequent to the point that some of the periodicals can be considered bilingual (Marcet i Salom 1987, vol. 2, 81). The earliest periodical publication related to the *Renaixença* movement in Catalonia was the fortnightly *Lo Verdader Català*, which only appeared from March to May of 1843 (six editions) and fell flat in a society still unfamiliar with this type of written Catalan (Marfany 2003, 649). Even its motto referred to the backward-looking view of the *Renaixença*: 'The prosperity of Catalonia and memories of its greatness'. Another early periodical is *Lo Gay Saber* (1868–1869 and 1878–1883) the title of which alludes directly to the cultural background of the *Jocs Florals*. It is important to stress that this fortnightly publication presented itself as 'made by Catalan, Majorcan and Valencian writers', a claim reflected, for example, in the collaboration of Joaquim Rubió i Ors from Barcelona, Teodor Llorente (1836–1911) from Valencia and Marià Aguiló (1825–1897) originally from Majorca.

As already mentioned, the main organ of the Catalan revival movement was *La Renaixensa*, founded in 1871 by Pere Aldavert (1850–1932), who served as its director and was later seconded by the playwright and poet Àngel Guimerà. The main ideas

published in this periodical revolved around Catalonia as a nation with the defining characteristics of its own language, history, law and economy and therefore in a position to restore its lost sovereignty (Duran i Tort 2001, 197). It should be highlighted that the periodical always strove to serve as a politically neutral platform for all those who felt strongly about Catalonia and the Catalan language. Only when *La Renaixença* became a daily newspaper (1881) did it adopt political positions and lend its support to the *Memorial de Greuges* ('Petition of Grievances', 1885), the *Missatge a la Reina Regent* ('Message to the Queen-Regent', 1888) and the 'Basis for a Catalan Regional Constitution' (1892), known as *Bases de Manresa* (cf. 5.1).

Still within the framework of political Catalanism, May 1879 saw the publication of the first daily newspaper written entirely in Catalan. Founded by the father of the Catalan left-wing nationalism Valentí Almirall (1841–1904), the *Diari Català* was in circulation until June 1881. The paper had to change its name several times throughout its lifetime due to the multiple repressive allegations commonly made against the progressive press during that period in Spain (Figueres 1999, 33–52). The newspaper attempted to raise awareness of issues specific to Catalonia and Catalans, particularly among the working classes, and was a vocal supporter of federal republicanism (Figueres 1999, 348–355). Nevertheless, it should be stressed once more that the political press – like most of the press printed in the Catalan-speaking areas – was in Spanish. In this regard, Almirall's engagement with federalism was visible in the biweekly journal *El Federalista* (October 1868–March 1869) and the newspaper *El Estado Catalán* (Barcelona, 1869–1872 and Madrid, 1873), both written in Spanish.

Special attention should also be paid to the very healthy satirical press in Catalan. These periodicals had a great impact among the working class and peasants, who were not used to reading the daily newspapers and got their information through picture-based publications, mainly read aloud. One of the most influential and long-running publications with a federalist standpoint was *La Campana de Gràcia* (1870–1934). It was initially produced with a print-run of just 3,000 copies, but circulation increased rapidly, soon averaging around 10,000. In 1888, it ran 20,000 copies and is thought to have reached 30,000 copies in 1903 (Capdevila 2014, 61). It was followed by *L'Esquella de la Torratxa* (1872, 1874, 1879–1939), which achieved, for example, a print-run of 25,000 copies in 1898 (Capdevila 2013, 34). This is particularly significant given that in 1900 Catalonia's entire population was less than two million. *L'Esquella* was a republican, freethinking publication that presented itself as 'a satirical, humorous, illustrated and literary periodical' (Capdevila 2013, 15). There was a considerable price gap between the two periodicals, with *L'Esquella* costing twice as much as *La Campana* in the late 1880s. While *La Campana* was more popular and belligerent and primarily written for the working classes, peasantry and a general public, *L'Esquella* had a moderate, more playful and less combative character aimed at a readership in a higher income bracket (Capdevila 2013, 32). It should also be mentioned that at the beginning of the 20th century both periodicals voluntarily adopted the orthographic norms of the IEC, which contributed to their popularisation (Capdevila 2013, 41).

In the other Catalan-speaking areas this type of press was also in rude health. In fact, it was in Valencia that the first satirical magazine in Catalan *El Mole* came out (1837). It was founded and directed by Josep Maria Bonilla (1808–1880) and showed a particular interest in the issues affecting farmhands and day labourers. It adopted an anti-centralist position and had an acute awareness of Valencian history. The use of Catalan in *El Mole* should be understood in pragmatic terms, subordinated to its proselytism: to instruct the people it was necessary to speak the language most widely understood, in this case, Valencian (Balaguer 1988, 76). This position influenced the written form and spelling solutions adopted in the magazine, which obeyed the principles of Valencian “tal i com es parla” (‘as people speak’; cf. 6.) and its closeness to the *apitxat* dialect (Balaguer 1988, 77; 78.1 Dialects).

In Palma (Majorca) a weekly satirical and humoristic magazine appeared in 1879. Called *La Ignorancia* (1879–1883), it continued in 1887 under the name *La Roqueta*. Contributors included the Majorcan writer and philologist Tomàs Forteza i Cortés (1838–1898) and the priest, folklorist, linguist and historian Antoni M. Alcover (1862–1932).

5 Language and nation building

5.1 Political Catalanism and language

After the failed attempt to redefine Catalonia’s place in the Spanish state during the six revolutionary years (1868–1874), the 1880s and 1890s saw the development of political Catalanism. In September 1868 the Glorious Revolution, a military uprising, supported by Spanish liberals and republicans, led to the dethronement and exile of Queen Isabel II. In 1869 a liberal constitution was enacted and on 11th February 1873 the First Republic was proclaimed. This short-lived political regime – it came to an end on 29th December 1874 – went through five presidents, among them the Catalan republican and federalist Francesc Pi i Margall (1824–1901).

These events obliged several political movements to critically revise their positions. It can be assumed that during the first phase of the *Renaixença* movement ‘Catalanism’ covered more or less a sense of devotion to Catalan language and history. In this second phase, however, the term gained a more political connotation as the belief that Catalonia’s political identity – or that of the Catalan-speaking areas – as a ‘poble’ (“Volk”) or nation should be recognised. Thus, the fruitless attempts to establish a federal state in Spain during the First Republic sowed the seed for the emergence of a new political doctrine, that of political Catalanism and, more precisely, Catalan nationalism with the aim of restoring self-government (cf. Rafanell 1997, 229–269).

Although Catalonia’s identity or otherness embraces different elements such as language, culture, history, economy and law, language as a distinguishing sign was

promoted as the most important. From the very beginning of the organisation of political Catalanism, language has been a defining feature, as can be traced from the discussions of the *Primer Congrés Catalanista* (1880), which brought together all those interested in ‘the progress and development’ of Catalonia (“Bases”, Figueres 1985, 67). Despite the discrepancies between the rather apolitical group including the editors and contributors of *La Renaixença*, and the political group of Almirall’s followers, the more than 700 participants agreed on the creation of an Academy of Catalan Language (*Acadèmia de la llengua catalana*). It was decided that the 25 members of this academy would be chosen from the *Mestres en Gay Saber* of the *Jocs Florals*. Based on the model of the French and Spanish Academies, this future Catalan academy was charged with fixing and polishing the language and compiling ‘a good dictionary and a good grammar’ (Figueres 1985, 146), in other words, ‘to watch over our dear Catalan language’ (Figueres 1985, 157). Nevertheless, it should be mentioned that a small minority made the case for other future options such as the adoption of a universal (though artificial) language or the adoption of one major existing language, implied to be Spanish.

As a result of this First Catalanist Congress, the *Centre Català* (Catalan Centre) was founded with the aim of defending Catalonia’s moral and material interests and achieving the union of all interested parties in that aim. Its first president was Frederic Soler (1839–1895), an outstanding playwright of Catalan Romantic theatre, although the driving force was Almirall. In 1883 the Centre organised the *Segon Congrés Catalanista* (‘Second Catalanist Congress’, 1883), which presented participants with a Programme of Catalanism (*Programa del catalanisme*). It is significant that the third point of this programme directly addressed the language: ‘The Catalan nationalists aim that: [...] THIRD. The Catalan language is declared an official language of Spain to the same extent as the other languages spoken in the nation’ (Illa i Munné 1983, 77).

On the initiative of the Catalan Centre, in 1885 the so-called *Memorial de Greuges* or *Memoria en defensa de los intereses morales y materiales de Cataluña* was handed to Alfonso XII of Spain (1874–1885). The complaints were formulated mainly by Almirall and a non-party commission and were adopted in a public ceremony chaired by Joaquim Rubió i Ors and in the presence of many figures in Catalan civil society. This first expression of political Catalanism requested that the Spanish government take a differentiated approach and implement a regional system in accordance with the existing Spanish regions. The predominant and ‘unifying nature’ of the Castilian people was contrasted with the ‘analytical character’ of the Catalans, a perspective that is also applied to the linguistic differences between Spanish and Catalan (1885, 40–41). Though the document mainly protested against the unification of Spanish civil law, at several points the Catalan language and its value to the Catalans was drawn upon to support their arguments. At the reception on 10th March 1885, the conservative politician Marià Maspons i Labrós (1840–1885) attested to the contemporary state of Catalan in his speech (Duran Solà 2009, 33): “No podemos usar nuestra lengua más que en nuestros hogares y en conversaciones familiares; desterrada de las

escuelas, lo ha sido más tarde de la contratación pública y también de los tribunales, en los cuales muchas veces, y por muy ilustrados que sean, ni los jueces entienden a los testigos y procesados, ni éstos entienden a los jueces” (Maspons i Labrós 1968, 31).

The amalgamation of forces which represented distinct political directions broke apart when the more conservative elements separated to form the *Lliga de Catalunya* (‘League of Catalonia’, 1887), among them Àngel Guimerà and the highly influential architect and politician Lluís Domènech i Montaner (1850–1923). The *Lliga* also took an active role in drafting the *Missatge a la Reina Regent*, which was brought to the Queen-Regent María Cristina (1885–1902), who was named queen of the *Jocs Florals* in 1888. The Queen-Regent visited Catalonia for the inauguration of the Barcelona Universal Exhibition (1888), for which the redesign of the former military garrison at the *Ciutadella* was completed. It is worth noting that the Queen-Regent was addressed in Catalan, a symbolic, but nonetheless remarkable act.

On the initiative of the *Unió Catalanista* (‘Catalanist Union’), a platform founded in 1891 to fight the Spanish civil code, a meeting in Manresa was organised. This resulted in the *Bases de Manresa* (1892) or *Bases per a la Constitució Regional Catalana* (‘Basis for a Catalan regional constitution’). The main points comprised home rule for Catalonia and the concession of administrative, economic and fiscal powers. One demand was that Catalan be made the sole official language: ‘BASIS 3rd. The Catalan language will be the only one that, with official character, will be able to be used in Catalonia and in the relations of this region with the central power’ (1900 [1892], 7).

After the decline of left-wing Catalanism under Almirall’s leadership, the conservative wing gained influence. Most of its ideas were formulated around the periodical *La Veu de Montserrat* (1878–1900). Published in Vic and founded by Jaume Collell (1846–1932), this publication was the organ of moderate Catalan Catholicism and could count among its collaborators of the extremely influential bishop Josep Torras i Bages (1846–1916). The Catalanism defended from its pages focused mainly on Catalonia’s religious identity and tradition, demanding the establishment of Catalan denominational schools and the use of Catalan in all religious services (Ramisa 1985, 87–92 and 161–163).

Perhaps the three most prominent articulations of Catalanism at the turn of the century were *Lo catalanisme* (1886) by Almirall, *La tradició catalana* (1892) by Torras i Bages and *La nacionalitat catalana* (1906) by Enric Prat de la Riba (1870–1917). In his book, Almirall reassessed and systematised the main ideas that can be identified in the left-wing manifestations of Catalanism. The subtitle of *Lo catalanisme: Motius que l’llegitiman, fonaments científics y solucions practicas* (‘Motives for its legitimisation, scientific bases and practical solutions’) provides a brief overview of the book’s content. The language issue is addressed as the most visible element of Catalan identity (2009, 98) and its use is defended in the following words: ‘We talk and write in Catalan, and we will not stop using it until such time as we have obtained the great reparations that are owed to us. The use of our language is the most eloquent demonstration of our personality and an unanswerable argument in favour of the

justice of our cause. As long as the Catalan language lives, every act of unification, carried out in any territory, will be a veritable act of tyranny' (2009, 90).

At the other end of the ideological spectrum was the right-wing Catalanism expressed in *La tradició catalana. Estudi del valor ètic i racional del regionalisme català* ('The Catalan Tradition. Study of the ethic and rational value of Catalan regionalism'), which was intended as a response to Almirall's *Lo catalanisme*. For a long time, it served as the breviary of Catalan Catholics, popularly summarised with the motto 'Catalonia will be Christian or will not be at all'. According to Duran Solà (2009, 45), in *La tradició catalana* Torras i Bages stresses that Catalanism has been more concerned about its external and visible expressions, and less about its spirit or soul, that is to say, the national spirit, which is the life-giving element that connects the institutions of the nation. The Catalan spirit, forged over the centuries, is essentially Christian and the Catalan language is its expression:

'[...] between the thought and its expression, that is, the language, there is an intimate relationship, like between a mould and the moulded one; [...]. Among all the social bonds, once Religion has been taken out, the language is what binds most strongly; it makes interaction closer, facilitates conversation and makes mutual relations warmer; therefore, it is the element that must not be forgotten by those who want to influence the people' (Torras y Bages 1913, 38–39).

More than one decade elapsed between these two contributions to political Catalanism from 1888 and 1892 and *La nacionalitat catalana* (1906). This period was characterised by two decisive events: the Hispano-American War (April–August 1898) and the subsequent loss of Cuba and the Philippines. Spain's new status as an empire without colonies also had consequences for Catalonia, particularly, the industrial bourgeoisie. This particular class was suddenly obliged to channel its political and economic potential into the creation of Catalanist parties and a joint programme of Catalanism.

In 1899 the *Centre Nacional Català* ('Catalan National Centre') was created to pursue the cultural and political aims set out in the *Bases de Manresa*. Two years later the right-wing *Lliga Regionalista de Catalunya* ('Regionalist League of Catalonia', 1901–1936) was formed, by merging the former militants of the *Unió Regionalista* ('Regionalist Union', 1899) and the *Centre Nacional Català*. It disseminated its ideology through the periodical *La Veu de Catalunya* (1899–1936), which started with a print-run of 3,000 copies (Figueres 2014, 166).

One of the most notable members of the *Lliga Regionalista* was the lawyer and politician Enric Prat de la Riba who argued that Catalonia was a nation with its own laws, language, art, collective spirit and thought (2007, 49 and 100). In line with German Romanticism and Johann Gottfried Herder's (1744–1803) idea of the *Volk* (Prat de la Riba 2007, 79–80), the language plays a crucial role in *La nacionalitat catalana*: 'The nation [poble] that has been unable to construct its own language, is a crippled nation, because the language is the most perfect demonstration of the national spirit and the most powerful instrument of nationalisation, and therefore of the conservation and life of the nationality' (Prat de la Riba 2007, 91)

By reviewing Catalonia's distant, but also recent past and that of the "païssos de llengua catalana" ('the Catalan-speaking countries'; 2007, 97–100), Prat de la Riba set out the principal ideas on the decay and restoration of Catalan, ideas which often persist to the present-day research on these topics despite sometimes being disguised as a new approach.

5.2 The construction of a symbolic repertoire

The last two decades of the 19th century and the first decade of the 20th century saw the establishment of the now generally acknowledged benchmarks of the Catalan community. As has been stated, language was defined as the most significant unifying element of this community. Another very important reference was law (cf. 2). However, as in many other cases, the concept of belonging to a group, particularly a nation, is transmitted through a set of elements such as the flag, the coat of arms, the national anthem or the national holiday.

Around 1890 the so-called *senyera*, that is to say, the flag of four red stripes on a yellow background, took root as the symbol of Catalanism and was accepted as the national flag in times by *Solidaritat Catalana* ('Catalan Solidarity', 1906), a broad-based Catalanist coalition (Anguera 2010c). The flag itself is based on the coat of arms of the Crown of Aragon.

The anthem *Els segadors* ('The reapers') is a song from the oral tradition and refers to the Catalan Revolt (*Guerra dels segadors*) from 1640–1652. It was performed for the first time in 1892 heard in the adaptation made by Francesc Alió (1862–1908). Its incorporation into the repertoire of the *Orfeó Català* (Catalan Choral Society), founded in 1891, was key to its dissemination (Anguera 2010a). In 1931 it was declared to be the national anthem of Catalonia.

Among other popular and well-established dances, in the second half of the 19th century the *sardana* circle dance became the 'symbol and mirror of the inner union of the people' (Anguera 2010b), as claimed in 1894 by the Catalan lawyer, politician and writer Terenci Thos i Codina (1841–1903).

Other elements from the religious tradition were also integrated in the symbolic repertoire of nation building. In 1881 Pope Leo XIII declared the *Mare de Déu de Montserrat* ('Our Lady of Montserrat') Catalonia's patron saint. She shares this status with *Sant Jordi* (Saint George), who gained in popularity throughout the *Renaixença* due to a more secular and combative patriotic component (Anguera 2010d, 7). In 1881 the *Associació Catalanista d'Excursions Científiques* ('Catalanist Association of Scientific Excursions') began holding an award ceremony on 23rd April (Saint George's Day) that turned into an annual ritual (Anguera 2010a, 33). Both this day, on which offerings of roses were made, and the 11th September were adopted as national days. According to Anguera (2008, 8), the marking of the 11th September started in 1886 with a commemoration of those who were killed defending the city during the siege of

Barcelona at the end of the War of the Spanish Succession (1702–1714). The commemoration took place at Santa Maria del Mar (Barcelona), the parish of the *Fossar de les Moreres* ('Cemetery of the Mulberry Trees') where the remains had been buried. For many years, the day was celebrated on a rather informal basis or even clandestinely. It was declared the *Diada* ('national holiday') in 1980 by the Catalan parliament.

Nothing comparable to these changes has occurred in Valencia or the Balearic Islands.

6 Language evolution and code conflicts

The evolution of Catalan during the 19th century is characterised by the further increase of exogenous interferences, particularly from the official national languages used in the different Catalan-speaking areas. Additionally, the lack of a reference standard encouraged the fragmentation of the traditional spelling into dialectally marked practices. The weight of the omnipresent foreign orthographies (Spanish, French) became more and more tangible. Furthermore, the urban-rural differences contributed to the widening gap between the, linguistically speaking, conservative local dialects in the countryside and the evolving urban forms of expression.

Nevertheless, in the second half of the 19th century the (historical) description of Catalan began to be approached scientifically. As a result, the myth of Catalano-Limousin unity (cf. Rafanell 1991) was invalidated and the identification of the Catalan dialects in terms of Limousin was deprived of its theoretical basis. Particularly notable contributions came from Manuel Milà i Fontanals (1818–1884), who also participated actively in the reestablishment of the *Jocs Florals*. In his work *De los trovadores en España* (1861) he became the first to identify the division of Catalan-speaking areas into two major dialectal blocks, Eastern Catalan and Western Catalan:

“Entre las muchas diferencias locales de pronunciacion en diversos puntos de Cataluña [...] se distinguen dos grandes divisiones: la parte occidental-meridional [...] en que se pronuncia el catalan con mas limpieza y en general como se escribe, y la parte oriental en que se altera la pronunciacion, en que las vocales son menos limpias y en que hay sustitucion de vocales. *Esta sustitucion es la de la a á la e y de la u á la o en todas las sílabas no acentuadas*” (Milà y Fontanals 1861, 462).

This and other pioneering elucidations deeply influenced the perception of the Catalan language among his contemporaries.

From the 1860s onwards, the need to standardise the language in order to achieve the full restoration of Catalan was also articulated. The use of Catalan in the press and in novels and theatre, beyond popular theatre, raised awareness about the spelling and grammatical problems originating from the lack of standardisation. It became imperative to define the reference model for literary use. Two main currents of opinion can be distinguished: the advocates of “català acadèmic” ('academic Catalan'), who

stood for a restoration based on the historical stages of the language, and the advocates of “català que ara es parla” (‘Catalan as people speak now’). Their differences arose in the solutions they stipulated for spelling, as can be seen in the controversies surrounding the values of the letter *x* and the feminine plural-ending *-as/-es* (Segarra 1985, 172–218), which also responds to diatopic patterns in line with Milà i Fontanals.

The current of academic Catalan can be subdivided into two further groups. On the one hand, there were those who wished to construct the literary language on the basis of Catalan as used until the middle of the 16th century (“català acadèmic de tradició antiga”) and, on the other, those who considered the Catalan from 1600–1800 to be the appropriate starting point for constructing the modern language (“català acadèmic de tradició moderna”; Segarra 1985, 218–224; Solà 1991, 102–103). The latter current can be traced back to Ballot (cf. 3) and included authors such as Antoni de Bofarull and the Hellenist Josep Balari i Jovany (1844–1904). Unlike this group, the advocates of the ancient tradition gave priority to archaisms and dialectal expressions that were assumed to be more authentic. Tomàs Forteza i Cortés (cf. 4.2) formulated the principles of this opinion in “Observaciones generales sobre la lengua materna”, written in 1886, but not published until 1915 (Forteza y Cortés 1915, 1*–27*).

Radically opposed to the current of academic Catalan, the advocates of ‘Catalan as people speak now’ pleaded in favour of the everyday, contemporary spoken language. This was consistent with the literary genres they dealt with, for example, comedies and farces, and the target audience, that is to say, the common people. The most famous exponent of this view was Frederic Soler (cf. 5.1), who went by the pseudonym Serafí Pitarra. However, the current had many other followers in each of the Catalan-speaking areas. The debate reached its peak between 1864 and 1874 (Segarra 1985, 231–243). Echoes of the debate can be traced in several grammar and spelling books as well as periodical articles of the time (cf. 4.2).

The idea of a literary language based on the spoken language was subject to heavy criticism. In their grammar Antoni de Bofarull and Adolf Blanch (1832–1887) put forward the ground-breaking argument: “No pretendemos [...] que sea la nuestra, gramática de la lengua catalana segun ahora se habla, pues sobre que en ningun tiempo y en ningun país se ha hablado generalmente como se ha escrito ó perorado, no nos hallamos en el caso de tomar mas del lenguaje comun que del literario, ya que de restauracion y de perfeccionamiento se trata” (Bofarull/Blanch 1867, 6). At the end of the century *L’Avenç* promoted the definition of a modern standard, with the first seminal attempt being Pompeu Fabra’s *Ensayo de gramática de catalán moderno* (2005, 1891) (14.1 The Language Reform, the Institut d’Estudis Catalans and the Work of Pompeu Fabra). Nevertheless, it seems that this codification followed certain trends in written use that were not yet mainstream but clearly becoming more widespread (Ginebra 2009, 332).

7 Bibliography

- Almirall, Valentí (2009, ¹⁸⁸⁶), *Lo catalanisme*. Presentació de Manuel Castellet i Marc Jamal, Barcelona, Centre Excursionista de Catalunya.
- Anguera, Pere (2008), *L'Onze de Setembre. Història de la Diada (1886–1936)*, Barcelona, Centre d'Història Contemporània de Catalunya.
- Anguera, Pere (2010a), *Els segadors. Com es crea un himne nacional*, Barcelona, Dalmau.
- Anguera, Pere (2010b), *La nacionalització de la sardana*, Barcelona, Dalmau.
- Anguera, Pere (2010c), *Les quatre barres. De bandera històrica a senyera nacional*, Barcelona, Dalmau.
- Anguera, Pere (2010d), *Sant Jordi, patró de Catalunya*, Barcelona, Dalmau.
- Balaguer, Enric (1988), *Una revista popular valenciana: "El Mòle" (1837 i 1840–41)*, Caplletra 4, 69–78.
- Ballot, Josep Pau (1987, ¹⁸¹⁴), *Gramatica y apología de la lengua catalana*, ed. Mila Segarra, Barcelona, Altafulla.
- Bases per a la Constitució Regional Catalana* (1892) = Unió Catalanista (1900), *Bases pera la Constitució Regional Catalana*, acordades per la Assamblea de Delegats celebrada á Manresa los días 25, 26 y 27 de Mars de 1892, Barcelona, La Renaixensa.
- Bofarull, Antoni de (1862), *La Orfaneta de Menargues ó Catalunya agonisant. Novela històrica*, Madrid/Barcelona, Llibreria Espanyola/Llibreria del Plus Ultra.
- Bofarull, Antoni de/Blanch, Adolf (1867), *Gramática de la lengua catalana*, Barcelona, Espasa.
- Brumme, Jenny (1997), *Spanische Sprache im 19. Jahrhundert. Sprachliches Wissen, Norm und Sprachveränderungen*, Münster, Nodus.
- Capdevila, Jaume (2013), "L'Esquella de la Torratxa". 60 anys d'història catalana, in: Jaume Capdevila (coord.), "L'Esquella de la Torratxa". 60 anys d'història catalana (1879–1939), Pròleg de Jaume Guillamet, Barcelona, Efadós, 13–58.
- Capdevila, Jaume (2014), "La Campana de Gràcia". La primera publicació catalana de gran abast, 1870–1934, Lleida, Pagès.
- Capmany y de Montpalau, Antonio (1779–1792), *Memorias históricas sobre marina, comercio y artes de la antigua ciudad de Barcelona*, 4 vol., Madrid, Antonio de Sancha.
- Decreto Romanones = Ministerio de Instrucción pública y Bellas Artes, *Real Decreto, Madrid 21 de Noviembre de 1902*, in: Gaceta de Madrid 327, 23 November 1902, 663–664.
- Domingo, Josep M. (2005), *Introducció*, in: Francisco M. Tubino, *Historia del Renacimiento literario contemporáneo en Cataluña, Baleares y Valencia*, edició facsímil en CD-Rom, Lleida, Punctum/Aula Màrius Torres.
- Domingo, Josep M. (2009), *Renaixença: el mot i la idea*, Anuari Verdaguer 17, 215–234.
- Duran i Tort, Carola (2001), "La Renaixensa", *primera empresa editorial catalana*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Duran Solà, Lluís (2009), *Breu història del catalanisme. I: Del segle XIX a la Dictadura de Primo de Rivera*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Fabra, Pompeu (2005), *Obres completes*, vol. 1: *Gramàtiques de 1891, 1898, 1912*, Barcelona, Institut d'Estudis Catalans.
- Ferrando Francés, Antoni/Nicolás Amorós, Miquel (2016), *Història de la llengua catalana* (Nova edició revisada i ampliada), Barcelona, UOC.
- Figueres, Josep M. (1985), *El Primer Congrés Catalanista i Valentí Almirall. Materials per a l'estudi dels orígens del catalanisme*, Barcelona, Generalitat de Catalunya.
- Figueres, Josep M. (1999), *El Primer diari en llengua catalana. "Diari català" (1879–1881)*, Barcelona, Institut d'Estudis Catalans.
- Figueres, Josep M. (2014), *La Veu de Catalunya (1899–1937)*, Barcelona, Ajuntament de Barcelona.

- Ford, Richard (³1855), *Handbook for Travellers in Spain. Part I. Andalucía, Ronda and Granada, Murcia, Valencia, and Catalonia*, Entirely revised, with great additions, London, Murray.
- Forteza y Cortés, Tomás (1915), *Gramática de la Lengua Catalana*, Obra premiada en el Certamen de Ferias y Fiestas de Palma de 1881. Con un prólogo por Antoni M.^a Alcover, Palma, Escuela Tipográfica Provincial.
- Ginebra, Jordi (2009), *La construcció de la llengua literària contemporània: què devem al segle XIX?*, *Anuari Verdguer* 17, 309–334.
- Hina, Horst (1978), *Kastilien und Katalonien in der Kulturdiskussion 1714–1939*, Tübingen, Niemeyer.
- Hina, Horst (1986), *Castilla y Cataluña en el debate cultural 1714–1939. Historia de las relaciones ideológicas catalano-castellanas*, trans. Ricard Wilshusen, Barcelona, Península.
- IEC = Institut d'Estudis Catalans (1992), *The Institut d'Estudis Catalans*, trans. Valerie Collins, Barcelona, Institut d'Estudis Catalans.
- Illa i Munné, M. Carme (1983), *El Segon Congrés Catalanista. Un congrés inacabat (1883–1983)*, Barcelona, Generalitat de Catalunya.
- Kailuweit, Rolf (1997), *Vom EIGENEN SPRECHEN. Eine Geschichte der spanisch-katalanischen Diglossie in Katalonien (1759–1859)*, Frankfurt am Main, Lang.
- Kailuweit, Rolf (1999), *El canvi de l'arquitectura lingüística de les terres catalanes en els segles XVIII i XIX*, *Caplletra* 27, 189–211.
- Kloss, Heinz (²1978), *Die Entwicklung neuer germanischer Kultursprachen seit 1800*, Düsseldorf, Pädagogischer Verlag Schwann.
- Laborde, Alexandre (1827), *Itinéraire descriptif de l'Espagne*. Tome second. Troisième édition, revue, corrigée et considérablement augmentée, Paris, Didot.
- Ley de Enjuiciamiento Civil* (1880), in: *Gaceta de Madrid* 53, 22 February 1881, 326–518.
- Ley del Notariado* (1862) = *Ley del Notariado de 28 de mayo de 1862*, in: Ministerio de Gracia y Justicia, *Boletín Oficial del Estado* 149, 29 May 1862.
- Ley del Registro Civil* (1870) = *Reglamento para la ejecución de las leyes de matrimonio y registro civil* (1870), *Gaceta de Madrid* 348, 14 December 1870, 1–4.
- Ley Moyano* (1857) = *Ley de Instrucción Pública*, in: *Gaceta de Madrid* 1,710, 10 September 1857, 1–3.
- Marcet i Salom, Pere (1987), *Història de la llengua catalana*, 2 vol., Barcelona, Teide.
- Marfany, Joan-Lluís (2001), *La llengua maltractada. El castellà i el català a Catalunya del segle XVI al segle XIX*, Barcelona, Empúries.
- Marfany, Joan Lluís (2003), *En pro d'una revisió radical de la Reinaxença*, in: *Professor Joaquim Molas: memòria, escriptura, història*, vol. 2, Universitat de Barcelona, 635–656.
- Maspons i Labrós, Marià (1968), *[Parlament davant el rei Alfons XII]*, in: Joaquim de Camps i Arboix, *El Memorial de Greuges*, Barcelona, Dalmau, 29–31.
- Memoria en defensa de los intereses morales y materiales de Cataluña* presentada directamente á S.M. el rey: en virtud de acuerdo tomado en la reunión celebrada en la Lonja de Barcelona, el día 11 de enero del año 1885, Barcelona, Imprenta Barcelonesa, 1885.
- Milà y Fontanals, Manuel (1861), *De los trovadores en España. Estudio de lengua y poesía provenzal*, Barcelona, Verdguer.
- Pers y Ramona, Magí (1847), *Gramática catalana-castellana*, Barcelona, Berdeguer.
- Pinyol i Torrents, Ramon (2009), *Jacint Verdguer, poeta nacional de Catalunya*, *Catalan Historical Review* 2, 213–226.
- Prat de la Riba, Enric (2007), *La nacionalitat catalana*, edició facsímil, Barcelona, Escola d'Administració Pública; (¹1906), Barcelona, L'Anuari de la Exportació.
- Rafanell, August (ed.) (1991), *Un nom per a la llengua. El concepte de llemosí en la història del català*, Vic, Eumo.
- Rafanell, August (1997), *El català al segle XIX. De llengua del poble a llengua nacional*, Barcelona, Empúries.

- Rafanell, August (1999), *La llengua silenciada. Una història del català, del Cinccents al Vuitcents*, Barcelona, Empúries.
- Ramisa, Maties (1985), *Els orígens del catalanisme conservador i "La Veu de Montserrat" 1878–1900*, Vic, Eumo.
- Schmid, Beatrice (2014), *Presencia y percepción del castellano en tratados de gramática y ortografía catalanas decimonónicas*, Boletín Hispánico Helvético 23, 227–245.
- Segarra, Mila (1985), *Història de l'ortografia catalana*, Barcelona, Empúries.
- Solà, Joan (1991), *Episodis d'història de la llengua catalana*, Barcelona, Empúries.
- Tayadella, Antònia (2012), *Novel·la i llengua al segle XIX: història i conflicte*, in: Antònia Tayadella, *Sobre literatura del segle XIX*, ed. Josep M. Domingo. Universitat de Barcelona, Societat Verdaguer, 97–109.
- Torras y Bages, Josep (³1913), *La tradició catalana: estudi del valor ètic i racional del regionalisme català*, Barcelona, Editorial Ibèrica.
- Tubino, Francisco M. (2005, ¹1880), *Historia del Renacimiento literario contemporáneo en Cataluña, Baleares y Valencia*, edició facsímil en CD-Rom, Lleida, Punctum/Aula Màrius Torres.

14 Towards Language Institutionalization

August Rafanell

14.1 The Language Reform, the Institut d'Estudis Catalans and the Work of Pompeu Fabra

Abstract: The Catalan language reached the end of the 19th century with a well-established *scripta*. The historical subordination of the Catalan language to Spanish or French (or Italian in the town of l'Alguer) contributed to the maintenance of a writing system distinct from the usual forms of the spoken language. Around 1890, Pompeu Fabra and his intellectual circle promoted the task of updating the Catalan corpus based on the most widely spoken dialect, supplemented by elements of the language from before the encroachment of the Spanish, French or Italian languages. The proposed reform was immediately accepted by the majority of the media, though it was above all driven by the most influential local institutions. Resistance to change, based on traditionalism or on dialectalism, ultimately failed. Therefore, Fabra's teachings ended up generalizing the idea of a flexible and modern "national language" with the arrival of the official status of the Catalan in the 1930s.

Keywords: Catalan language, Pompeu Fabra, anti-Fabrism, linguistic reform, language institutionalization

1 The inherited language

In 1815, Josep Pau Ballot published the first printed grammar specifically devoted to the Catalan language. Until then, the codification of Catalan had largely taken place on a social level by means of school textbooks, especially those for the purpose of learning Latin. The existence of such guides explains the continuous presence of Catalan on the book market, which indicates that, despite the political and cultural constraints imposed throughout the Catalan-speaking territories in the modern age, the publication and distribution of works in the local language was never brought to a complete halt. Ever since the arrival of the printing-press in the Catalan-speaking countries, hardly a single year has passed without the publication of books of one genre or another in Catalan (Lluch 1995).

Such perseverance helps to explain, to a large extent, why Catalan could not be reduced to the category of a non-standard language and Ballot himself said that it was

<https://doi.org/10.1515/9783110450408-021>

proof positive that Catalan should not be mistaken for some miserable dialect of Southern France. The steady stream of books published in Catalan led him to choose to base his model for the language he was attempting to standardize on precisely this continuity with the past which, despite the pressure from more powerful competitors, has never been interrupted. For Ballot, the relative isolation of the written standard from the living dialects did not present a problem because the broad acceptance of Catalan as a fixed language was taken as given.

As the diglossia of Catalan-Spanish (and Catalan-French north of the Pyrenees) continued to grow stronger in the Catalan-speaking territories, an established model for the language was seen as a buffer against the process of language shift. Disciples of Ballot soon followed, such as Magí Pers i Ramona, with his *Gramàtica catalana-castellana, adornada ab exemples de bons autors* (1847), and Pau Estorch i Siqués, with his *Gramàtica de la lengua catalana* (1857). A dignified and refined Catalan did still exist, if only as a symbolic compensation for its subordinate position. In Catalonia and in the Balearic Islands, the various men of letters who, at the beginning of the 19th century, felt the need to write in the local language knew how to do so because tradition was on their side. Even in the part of Catalonia administered by France, where the local language was more severely vilified than anywhere else in the Catalan-speaking communities, a grammar was published in 1852 that confirmed this general trend. Its author, Pere Puiggarí, declared his purpose as “le désir d’arrêter, autant que possible, la corruption qui dégrade de plus en plus la langue catalane, particulièrement en Roussillon” (1852, 2). Nevertheless, the grammatical conservatism of Ballot and his followers inevitably clashed with what sprang forth in the conversations of ordinary people. This disconnection between speech and writing was pointed out by the likes of Antoni Puigblanch or Joan Petit i Aguilar. Their criticisms, however, remained unpublished unlike the *Gramàtica de la lengua mallorquina* by Joan Josep Amengual (1835), in which real use of language emerges to challenge the established conventions.

Along with the prescriptive grammars, mid-19th century Catalan boasted a number of dictionaries certifying its status as an independent language. The *Diccionario catalán-castellano-latino*, issued in two volumes in 1803 and 1805, was compiled by various authors and organized by the Reial Acadèmia de Bones Lletres de Barcelona over the preceding decades. The words selected for this great work included not only those in contemporary use, but also those that shed light on the medieval and modern development of the language. However, the work that truly paved the way for Catalan lexicography until well into the 20th century was the *Diccionari de la llengua catalana ab la correspondencia castellana y llatina* by Pere Labèrnia (1839–1840), of which various editions were published throughout the rest of the century. Originally from Valencia but trained in Barcelona, Labèrnia lamented in the introduction to his work the ‘abject state of neglect that [Catalan] is in today, altered and disfigured by the mixture of foreign terms, idioms and accents’ (“estat de abjecció y descuyt com se troba en lo dia, alterat y desfigurat per la mescla de termes estranys, locucions y

accents"). For this reason, Labèrnia endeavoured to 'find the lustre of our native language in other periods than ours, in the monuments of our old literary glories' ("buscar lo llustre de nostra llengua nativa en altràs èpocas que la nostra, en los monuments de nostras antigüas glorias literarias", 1839, vol. 1, 3).

All of these works emerged at a time when Catalan was undergoing an accelerated process of diglossic subordination. The structures of the central government, set up in the early 18th century, were consolidated by the liberal regimes of the first half of the 19th century. In the mid-1800s, it was not only in French Catalonia that the historic language was restricted to unofficial use. It was also dying out as a means of higher expression in the rest of the Catalan-speaking territories. A number of educational and administrative laws served only to accelerate this trend. Thus, at this precise moment in its history, conditions were ripe for Catalan to be considered (as it often still is from positions of power) little more than a "dialecto regional".

It was around the 1860s that Catalan began to revive itself culturally, reflecting the memory of a lost golden age as well as a spontaneous surge of popular feeling that has never gone away. The seeds planted by Romanticism in the Catalan-speaking countries during the 1830s led inexorably to the idealization of the ancestral language. The change that took place among the country's elite would take on an air of restitution. The reconvening, in 1859, of the *Jocs Florals* of Barcelona – traditional poetry competitions, all but forgotten during the modern age – was an early sign of civil organization aimed at celebrating the language that represented the Catalans' heritage. The initial unfamiliarity caused by Catalan being elevated to a condition that had previously only belonged to the 'national language' ("llengua nacional") would gradually fade. With the rituals, speeches and publications of the *Jocs Florals*, Catalan recovered an arena exclusive to itself which, no matter how small, was an important step forward.

All of this immediately raised the question of what kind of Catalan was the most suitable for the output required by this new institutionalization of the language. In 1862, Antoni de Bofarull and Adolf Blanch, two of the original instigators of the *Jocs Florals*, published a *Gramàtica de la llengua catalana* which they hoped would become official within the atmosphere of this *Renaixença* (the name given to the renaissance of Catalan culture, ➤13 *Renaixença*). In reality, their book served to enshrine, with slight alterations, the work of Josep Pau Ballot. Traditional literary Catalan, conceived to fulfil functions limited to an imaginary recreation of reality, prevailed as the most suitable. It was exactly that: "català literari". For this reason, this model had to abstain from addressing both geographical and temporal concerns.

Initially, following the staging of the *Jocs Florals* in Barcelona, the debate on the characteristics of Catalan grammaticography had no significant impact. If anything, it merely focused on formalizing the spelling of the language. No solutions were agreed upon, leading to a general perception of "anarquia" well into the 20th century. To resolve the problem of alternative spellings, a proposal was made at the first Congrès Catalanista in 1880 to create an academy that would unify the language. But the

ideological heterogeneity of this politico-cultural movement meant agreement was impossible. While the more progressive contingent, led by the republican Valentí Almirall, proposed a modernization of Catalan orthography, the defenders of the *Jocs Florals* held fast to the preservation of fossilized conventions. From among this second faction, Francesc Matheu concluded his attack on reformist aims by requesting that ‘everyone write as they wish’ (“que cada hu escriga com vulga”), based on the argument that ‘uniformity would, today, be poverty’ (“l’unifomitat seria avuy probresa”, Matheu 1880, 126). The schism between those who saw tradition as untouchable and those who advocated modernization was irreconcilable. The impracticability of agreement on spelling would soon become patent, and the Acadèmia de la Llengua Catalana became inactive.

From the sidelines of the academic struggle, the critic Josep Yxart was able to comment, in about 1890, that as far as spelling was concerned, “cada escritor tiene la suya, y un criterio en que basarla”. In contrast to most of his contemporaries, Yxart felt there was no need for any regulatory body: “que un autor, un filólogo, plantee una ortografía uniforme, la base en un método racional, y él se impondrá a la larga; de la discusión brotará la luz; a despecho de la resistencia de algunos criterios particulares, el uso corriente y más común hará lo demás”. Without any fixed rules, the Catalan language had remained “intacta, inteligible y clara para millares de catalanes”. People in La Jonquera and those in Tortosa were still able to understand one another because “un idioma hablado persiste siglos por encima de las corruptelas de sus propios escritores, y hasta de la oposición de los que lo hablan” (Yxart 1890, 76–77).

Meanwhile, in French sovereign territory north of La Jonquera, the Catalan language remained active among ordinary people but was still scarcely found in writing. The influence of the *Renaixença* had arrived belatedly, second-hand from the south. The model for the very few printed texts from the Rosselló area relied on traditional language, maintained especially in ecclesiastical circles (religious sermons and texts). But of course, at the end the 19th century, modernity had come to Rosselló exclusively in French, above all as a result of legislation on free and compulsory education introduced by the Third Republic (1881–1882). Consequently, the transcription of the local language would be based on the established code of the dominant language. In 1894, the Perpignan poet and playwright Albert Saisset published a *Grammaire catalane* which, in practice, entailed a complete capitulation of written Catalan to the omnipotence of the French language.

A similar phenomenon was underway in Valencia, where the more popular publications rendered the local dialect according to Castilian orthography. The *Diccionario valenciano-castellano*, by Josep Escrig (1851), helped put a stop to this tendency. Revised in 1887, with the *Renaixença* in full flow, the Valencian language established in its pages looked to reconcile itself with older usage. Constantí Llobart, the curator of this third edition, conceded that there were three ways to read and write his language: one was archaic; the second was modern and more or less purged of “las ajenas influencias que hoy lo afean” and, finally, there was “el que en el día se

gasta" (Llombart ³1887, XVII–XVIII). An illustration of this is Josep Nebot's *Apuntes para una gramática valenciana popular* (1894), guided by natural usage but with Spanish spelling, which was followed by his *Tratado de ortografía valenciana clásica* (1910), in which he presented a counterargument to his previous work with archaic spellings.

On paper, the dichotomy in Rosselló and Valencia hardly differed from the one described in Catalonia by the renowned academic Josep Balari i Jovany. In his speech at the *Jocs Florals* of 1894, Balari noted the existence of four varieties of Catalan writing: one was medievalized, one was more modern; then there was the so-called 'academic' variety ("acadèmica", a mixture of the first two) and, finally, the one based on spoken Catalan. The orthographic misunderstanding was no more than a secondary symptom of the diglossia, which left room for only one undisputed, important language: the official one.

2 The ideological basis of the Reform

In the middle of this debate on orthography, Valentí Almirall began publishing the *Diari Català*, the first modern newspaper written exclusively in Catalan. In the prospectus announcing its arrival, it was clear that this new publication was not a conventional product. The *Diari Català* had to prove its progressive credentials at all levels, including in terms of the language itself. The aim, therefore, was to distinguish itself from the somewhat rigid language used in other Catalan printed texts, which largely consisted of literary-themed journals. Almirall's newspaper also distanced itself from the proponents of vulgarism, such as the weekly *La Campana de Gràcia* (1870–1934) or *L'Esquella de la Torratxa* (1872–1939), which enjoyed a wide circulation and popularity among the working classes.

The *Diari Català* (1879–1881) did not survive long, but it planted a fertile seed in the mentality of the time. Upon its closure, in 1881, Jaume Massó i Torrents, a disciple of Almirall, writing in *Lo Velógrafo* (1881), endorsed a linguistic Jacobinism, declaring, in the initial proposal, that they would not publish 'any work in the Majorcan or Valencian dialect, demanding of all the use of the true language, with the same orthography as we have adopted' ("cap treball en dialecte mallorquí ni valencià, exigint-se tots en el vertader idioma i amb la mateixa ortografia que hem adoptat", Miracle 1968, 220). Who knows whether what lay behind this option was the first description of contemporary Catalan established by the scholar Manuel Milà i Fontanals in 1875 (Milà y Fontanals 1890). In any case, in the same year, 1881, Massó founded a monthly magazine called *L'Avenç*, and at the presentation he declared that, unlike those who sounded the retreat, this new periodical would be guided by Progress. Nevertheless, it would not be until the second stage of *L'Avenç* (1889–1893) that Massó and his colleagues would openly consider the need to apply their concept of progressiveness to Catalan writing.

In 1890, Eudald Canivell, a typographer imbued with libertarian ideas, began the campaign in favor of a review of literary Catalan, of a Catalan that was no longer only confined to literature, nor to writing *teatre de costums* (“Costumbrismo”), which typically attempted to reproduce, more or less faithfully, the voice of ordinary people. Catalan also had to be able to describe the ideas of the present and project them into the future. For the purposes of the new Catalanist tendency, the conventionalities of the old grammar and spelling rules were nothing but old customs. According to Canivell (1890), the language had to be ‘correct while, at the same time, modernized’ (“correcte al mateix temps que modernisat”). The standard language had to be established by ‘bringing into agreement the written language and the spoken language, knowing how to separate the pure from the flawed’ (“posant d’acort lo llenguatge escrit ab lo llenguatge parlat, sabent separar en aquest lo pur de lo viciat”, 1890, 157). For Canivell, the choice of one dialect or another was not paramount. What was paramount was that while there was no such agreement, efforts had to be made to ‘resist foreign invasion, especially from Spanish’ (“resistir la invasió forastera, sobre tot la castellana”), even though this may mean ‘fraternizing with solutions from the French or Italian languages’ (“agermanarse á la francesa o italiana”, 1890, 162).

The article by Canivell led to a press campaign on the need to reform written Catalan. Soon afterwards, *L’Avenç* became *L’Avenç*, in an explicit attempt to recover the classic Catalan grapheme, which had been abandoned over the previous centuries, coinciding with French and simultaneously diverging from modern Spanish. From now on, the instigators of *L’Avenç*, would take on this reforming mission. According to one of them, Pompeu Fabra, born in the Vila de Gràcia, Barcelona, ‘historical grammar’ was one thing, while prescriptive rules for the present were quite another. To that end, ‘the current forms and pronunciation of Eastern Catalan should be definitively adopted due to their undeniable superiority’ (“les formes actuals i la pronunciació oriental han de ser definitivament adoptades per la seva innegable superioritat”, *LV*, 30-XII-1891). In 1891, Fabra published his *Ensayo de gramática de catalán moderno*, the main purpose of which was ‘to present Catalan as it is spoken’ (“ha sigut presentar el català tal com se parla”) on the understanding that the Catalan used in the Barcelona area was the same that was used ‘with small variations, in the whole of Catalonia, except in some areas in the provinces of Lleida and Tarragona’ (“am petitas modificacions, en tot Catalunya, excepte en algunas comarcas de les provincias de Lleyda y Tarragona”, Fabra 2005, 209). The reformed Catalan language had to be based on the most widely spoken variety of Catalan, which was also the most phonetically different from the dominant Spanish. At a conference at the *Lliga de Catalunya*, Fabra again declared bluntly: ‘We have preferred the dialect that is most antagonistic to Spanish; we have endeavored to prove its goodness’ (“Hem preferit el dialecte més antagonic de l’espanyol, ens hem esforçat a provar la seva bondat”, Casas-Carbó 1892, 223).

The change of perspective wrought by the *L’Avenç* group was absolute – to the immense discomfort of the traditionalists, who saw these modifications as juvenile provocation. One of them referred to the reform proposals in terms of a “verdadera

revolución en nuestro lenguaje escrito” (LV, 25-XII-1891). Joaquim Casas-Carbó, a partner of Massó i Torrents in the publishers of *L'Avenç*, agreed entirely: “Se ha dicho que con éstas [reformas] introducimos una perturbación en el campo de las letras catalanas; que pretendemos hacer una verdadera revolución en el lenguaje literario. El que tal dice está en lo cierto. Pero como nosotros demostramos que esta revolución es justa y necesaria, en vez de rechazar el dictado de revolucionarios (lingüísticamente hablando por supuesto) como una acusación, lo aceptamos con júbilo” (LV, 30-XII-1891). Under the stewardship of Massó i Torrents and Joaquim Casas-Carbó, the publishers of *L'Avenç* began putting these reformist theories in print. *Montalba*, the novel by Carles Bosch de la Trinxeria – an author of traditionalist ideology – was the first literary work to be modified according to these theories. Canon Jaume Collell, promoter of the conservative clerical weekly *La Veu de Catalunya*, compared it to expropriation: “They have castrated the most manly and sturdy of the Neo-Latin languages, giving it the air of a Tagalog dialect or some such thing” (“Han castrat l'idioma més viril i ferreny dels neo-llatins, donant-li els aires d'un dialecte *tagalo* o cosa així”, Requesens i Piqué 1994, 90). What scandalized the conservatives was not only the restitution of the letter ç, but also a plethora of forms – graphemes, phonemes or grammar – that in many cases lacked historical documentation, and which were now to be taken as standard.

In February 1891, *L'Avenç* changed the title of a story by Joan Pons i Massaveu from *Nadal a Brugarolas* to *Nadal a Brugaroles*, thus aiming to put an end to the controversy over the spelling of the plural forms of nouns and verbs ending in *a* or *e*, which had dominated much of the debate during the *Renaixença*. As Casas-Carbó said in the most widely read daily newspaper in Catalonia (published in Spanish, of course), conventional Catalan had separated itself ‘unjustifiably’ (“injustificadament”) from spoken Catalan, using “convenciones ortográficas absurdas e impropias de una lengua moderna”. It was now time to resolve this divorce using “razones sistematizadas” (LV, 25-XII-1891).

3 The launch platforms

The turn of the century in Catalonia was characterized by a proliferation of nationalist discourse and the defining character of the “*nació catalana*” was, above all, the existence of its own, distinct language. Accordingly, the early waves of political nationalism went hand in hand with an emerging linguistic nationalism that claimed that the territorial reach of the language corresponded with that of the nation. The *floralists* had published the first map of the Catalan language in the magazine *La Il·lustració Catalana* in 1883. It was not by accident, then, that in 1900, this map was adopted by the founders of *L'Avenç*, whose platform was now the journal *Catalonia*. However, for them, the demarcation of ‘the Catalan lands’ (“*les terres catalanes*”) was more than a linguistic map (Rafanell 2011, 98–99); the concept of “*Catalònia*” in-

cluded areas to the north of Perpinyà down to the south of Alacant and from the west of Lleida to the Balearic Islands, a geographical demarcation whose destiny was to become a national (i.e. political) territory. In addition, the pages of *Catalonia* would also consolidate the orthographic and grammatical reforms which were continuing to be developed and promoted with publications from the publishing house of *L'Avenç* (which survived the demise of the journal), in particular, through the pocket-sized editions of its famous *Biblioteca Popular* (1905–1926), which produced works by both Catalan and foreign authors.

1899 saw the first issue of *La Veu de Catalunya*, a daily newspaper covering a broad spectrum of the blossoming Catalan nationalism which, from the start, took into account the full sweep of the Catalan-speaking territories. That said, in the early years of the 20th century, the greatest proponent of the idea of Catalan linguistic unity was a more peripheral figure. In 1901, Mossèn (“priest”) Antoni M. Alcover, vicar general of Majorca, born and raised on the island in the town of Manacor, sent an open letter to the Catalan faithful asking for their support in a colossal task:

‘We want to make an inventory, everything that we know, of the vast richness, the imponderable opulence, that our language possesses in words, sentences, adages, idioms and forms, spread far and wide in the numerous written monuments of her children since the eleventh century, kept safe in libraries and archives: a richness and opulence that still springs forth, in torrents, flashing by, unstoppable, from the mouths of the thousands and thousands of people who inhabit Spanish Catalonia and French Catalonia, the Balearic Islands and the ancient kingdom of Valencia’.

“Volem fer l’inventari, tot lo complet que sabrem, de la riquesa, de la opulència imponderable, estupenda, que en paraules, frases, adagis, modismes y formes té escampada y espargida la nostra llengua dins els nombrosos monuments escrits dels seus fills del segle XI ensà, guardats dins biblioteques y arxius: riquesa y opulència qui brollen encara rabents, llampants, inestroncables, de la boca dels milenars de milenars de gent qui pobla Catalunya espanyola y Catalunya francesa, les Illes Balears y l’antich regne de Valencia” (Alcover 1901, 7).

Guided by the lexicographical work undertaken by his fellow Majorcan, Marian Aguiló (1822–1897), Alcover used his status as a man of the Church to encourage the commitment of a dense network of collaborators. Over time, Alcover’s dictionary project came to comprise nearly one million index cards, collected with the aim of constructing an entire inventory of Catalan: the ancient and the modern, the dialectal and the literary. To keep people informed of his ongoing project, Alcover periodically published the *Bolletí del Diccionari de la Llengua Catalana* – the first Hispanic philological magazine in history – which became a platform for articles and book reviews on Catalan, reports of Alcover’s philological excursions up and down the country or the monographs defining a particular aspect of the language. In early 1903, “Cataluña bilingüe”, an article by Ramón Menéndez Pidal was published. Alcover used the *Bolletí* to express his contempt for it in his “Qüestions de llengua y literatura catalana” – the most extensive and most widely disseminated history of the Catalan language (very *sui generis*, it must be said) that there had ever been.

In January 1904, Alcover promoted the idea of a conference on Catalan grammar. The poet Josep Carner declared himself fascinated by the courage of this newcomer from Majorca: 'Other writers have laboriously carved out isolated rivers. Mossèn Alcover has chosen the sea where, sooner or later, all the rivers arrive' ("Els altres escriptors han format treballosament els rius isolats. Mossèn Alcover ha escollit el mar, on arriben tard o d'hora els rius", *LVC*, 26-VIII-1904). The conference was postponed as a result of the events of November 1905, when members of the Spanish military raided and destroyed the offices of the weekly satirical magazine *¡Cu-Cut!* and the newspaper *La Veu de Catalunya*. Once these two Catalanist periodicals were permitted to publish again, they adopted some of the practices of *L'Avenç* and *Catalonia*, such as the emblematic representation of plurals ending in *-es*. The progressive newspaper *El Poble Català* had already adopted such practices since its inception in 1904 (Rafanell 2011, 165–180). Finally, armed with the broader aim of studying all aspects of the language, the First International Congress of the Catalan Language was held in Barcelona in October 1906. A few months earlier, Enric Prat de la Riba, editor of *La Veu de Catalunya*, had written *La nacionalitat catalana*, an essay which immediately became essential reading for the first wave of Catalan nationalism. Prat de la Riba described the Catalan language as a basic ingredient of ethnicity and as 'a natural product, not the result of some convention or artifice of man' ("producte natural, no el resultat d'una convenció o de l'artifici de l'home", Prat de la Riba 2000, 148).

Among the more than 3,000 attendees at the Congress were all the major Catalan literary names of the time, with virtually no exceptions. Some of the world's most prestigious Romance scholars also attended, including Menéndez y Pelayo, Fastenrath, Saroïhandy, Guarnerio, Foulché-Delbosc and Vogel. The composition of the technical committee of the congress reflected an express desire to settle old disputes. With Alcover as chair, there was Antoni Rubió i Lluch (professor of Catalan literature in the recently established, and unofficial, *Estudis Universitaris Catalans*) along with Jaume Massó i Torrents and Joaquim Casas-Carbó, founding members of *L'Avenç*. In his speech, Prat de la Riba stated that 'language is nationality itself' ("la llengua és la mateixa nacionalitat"). However, the real scientific revelation of the Congress came from the contribution from Pompeu Fabra, then a teacher of engineering in Bilbao. Fabra presented an important paper addressing several burning issues on Catalan spelling. He particularly excelled as he dismantled just about everything Alcover had said in his talk. Fabra's criticism was made from a standpoint of syntactic dissection based on a comparison with other languages and, first and foremost, 'the current state of the language' ("l'estat actual de la llengua", Fabra 1980, 79–80), by which he meant, primarily, the state of the language in the central area of Catalonia.

Joan Maragall, the most widely recognized Catalan writer at the beginning of the century, also took part in the Congress. A former collaborator of *Catalonia* and friend of the modernists who recognized a focal point for collective regeneration in the Catalanist movement, Maragall delivered a fierce defense of dialectal freedom above

and beyond the arbitration of grammarians. In his eyes, something as alive as a language of the people was not 'legislable' ("llegislable"). The imposition of any one variety over another 'could distort [...] the spiritual emergence of Catalonia, corrupting its highest verbal expression' ("vindria a falsejar [...] l'arrencada espiritual de Catalunya, corrompent la seva més alta expressió verbal". CILC 1908, 490). Maragall warned the Catalanists not to let themselves be fooled 'by an external, arbitrary ideal of false unity' ("per un ideal de falsa unitat externa", CILC 1908, 493). These sentiments from its most eminent member were echoed in the conclusions of the Congress: 'Catalan literature must not only maintain and use, but also promote and dignify all the dialectal variations that are the lifeblood of the language and, therefore, of literature itself' ("La literatura catalana ha de, no sols mantenir y utilitzar, sinó fomentar y dignificar totes les varietats dialectals que són la vida de la llengua y, per tant, de la literatura mateixa", CILC 1908, 688). But in 1908, when Maragall's paper was published, the proponents of modern literary language were already challenging this idea: 'We have to conclude that the *word* has never been alive; but that maybe one day it will become so, when, being so perfect, all its elements take on, around a nucleus, the supreme form of an exquisite polygon' ("Devem pensar que la Paraula no ha estat mai viva; però que acàs un jorn ho devingui, quan, de tan perfecta, ja tots els seus elements prenguin, entorn d'un núcli, la forma suprema d'un esquisidíssim polígon", Ors 1908, X).

One year after the Congress, Pompeu Fabra made his thoughts known on various outstanding problems in the article "Sobre diferents problemes pendents en l'actual català literari". The juncture he proposed between the living word and the standardized language also seemed to be an answer to the dialectal hypersensitivity of Maragall:

'In fact, if every writer wrote in *his own* Catalan – not the Catalan of the inconsequential people of his environs but the Catalan of the most educated people – the divergences that might appear in the literary language would perhaps not be, nowadays, as great as many think, and it seems true that, as favorable times for Catalan nationality are coming, these divergences will dwindle further, and there will be an on-going rapport between the literary language and the language of the capital: not the Barcelona dialect of today, but rather, a Barcelona dialect free of Castilianisms, one that is influenced by all the other Catalan dialects, enriched, refined: the future dialect of the future capital of Catalonia'

"En realitat, si cada escriptor escrivia en el *seu* català – no en el català de la gent ínfima de la seva encontrada, sinó en el de la gent més culta – les divergències que apareixerien en la llengua literària no foren potser, ja avui dia, tant grosses com molts pensen, i es de creure, venint temps favorables a la nacionalitat catalana, que aquestes divergències anirien atenuant-se, i una compenetració s'aniria operant entre la llengua literària i el llenguatge de la capital – no'l barceloní d'avui dia, sinó un barceloní tot altre, deslliurat de castellanismes, influenciat per tots els altres parlars catalans, enriquit, refinat: el futur parlar de la futura capital de *Catalonia*!" Fabra, Pompeu (1907, 369; punctuation has been slightly modified).

This text bore the seal of a brand-new institution. In 1907, as president of the Barcelona Provincial Council, Enric Prat de la Riba created the Institut d'Estudis

Catalans (IEC), the body charged with overseeing Catalan high culture. This had, without doubt, been made possible because 'the Catalanist movement had become a genuine mass movement and had been able to conquer certain spheres of political power' ("perquè el moviment catalanista havia esdevingut una realitat de masses i havia estat capaç de conquerir espais de poder polític", Balcells/Pujol 2002, 13). Prat divided the new institution into sections, one of which was in charge of studying and standardizing the historical language of the Catalans. In 1911, at the inauguration of the Secció Filològica ('Philological Section') of the IEC, Prat expressed himself in terms that could very well have been Fabra's. He pressed hard for an academy of philological research, because only the consolidation of one type of written language would prevent them 'falling into the grave danger of dialectal anarchy' ("guarda de caure en el perill gravíssim de l'anarquia dialectal"). The peoples left behind by history had not *one* literary language but as many as they had counties or regions. The same went for grammar, spelling or the lexicon. In medieval times, unity was the norm. But in the case of Catalan, that normal evolution had been stifled in the early centuries of the modern age. The Secció Filològica could now make up for lost time:

'The mission, first and foremost, is to discover and formulate the grammatical laws of our language, ascertain and standardize the spelling, compile the complete and whole Catalan lexicon, with the etymology, phonetic notation, all the dialectal varieties, the historical evolution of the words, the geographical area of each one, the graphic expression of the things that each word has represented or represents and its scientific definition.

All appropriate means for such a mission must be employed by this Institute; historical research; geographic research; field work; phonetic studies; comparative studies; consultations; public information; public competition; organizing collaboration in all the lands of the Catalan language, from Perpinyà to Elx and from Fraga in Aragon to l'Alguer in Sardinia; the cooperation of expert individuals in the different strata of social activity; discussions with the whole *Institut d'Estudis Catalans*; publication of texts and questionnaires and instructions for those cooperating with partial or occasional studies'.

"La seva missió primera, primordial, és descobrir i formular les lleis gramaticals del nostre idioma, escatir i fixar les seves formes ortogràfiques, inventariar el lèxic català, totalment, integralment, amb filiació etimològica, amb notació fonètica, amb totes les varietats dialectals, amb l'evolució històrica dels mots amb l'àrea geogràfica de cada un, amb l'expressió gràfica de les coses que cada paraula ha representat o representa, amb la seva definició científica.

Tots els mitjans adequats a una semblant missió han de ser usats per aquest Institut; investigacions històriques; investigacions geogràfiques; excursions; estudis fonètics; estudis comparatius; consultes; informacions públiques; concursos; organització de la col·laboració en totes les terres de llengua catalana, des de Perpinyà a Elx i des de Fraga d'Aragó a l'Alguer de Sardenya; cooperació d'elements tècnics en els diferents ordres de l'activitat social; discussions amb l'Institut d'Estudis Catalans en ple; publicació de textos i de qüestionaris d'instruccions per als cooperadors d'estudis parcials o fragmentaris" (Prat de la Riba 2000, 580–581).

Now was the time to activate what Eugeni d'Ors, secretary general of the IEC and popular writer of dogmatic columns in the press, described as the 'holy work of the philological police' ("santes empreses de filològica policia", *LVC*, 17-I-1912). In 1912, the Secció Filològica, presided over by Alcover, delegated one of its members, Pom-

peu Fabra, to direct the *Oficines Lexicogràfiques*, tasked with preparing the future general dictionary of Catalan and editing materials from the legacy of Marian Aguiló. In 1912, the Barcelona Provincial Council granted Fabra the chair of Catalan at the Universitat Industrial. That same year, having moved to Catalonia once and for all, Fabra published, in Spanish, his *Gramática de la lengua catalana*. In addition to presenting Catalan as a language comparable to any other, with a broad geographical reach and solid credentials, Fabra describes its modern features with additional information on the main dialectal alternatives and the archaic morphological variants. It was a seminal moment in the institutionalization of the language reform conceived twenty years previously by a group of perfectly ‘revolutionary’ individuals.

4 The major works of the new order

In the same year the Secció Filològica of the IEC was created, a discussion paper was drawn up to establish Catalan spelling and grammar rules. Most of this work was carried out by Pompeu Fabra. After lengthy discussions over the details, the *Normes ortogràfiques* were approved by the board of the IEC and made public in January 1913. From the pages of *La Veu de Catalunya*, Prat de la Riba presented the result as a milestone for the country. This “naturalist” politician had surrendered to the decisions of the experts in order to preserve the unity and uniqueness of the language:

‘Creating a literary language is the crowning work of a people [...]. When, from the whole array of a language’s dialects, one stands out and then soars above the others, giving shape to the literary and philosophical and legal creations of the race, that language – in its natural state we spoke of, that incoherent throng of dialectal speech – becomes articulated and becomes a living unit in a center of reconnection and coordination and purification of all the linguistic forces of the race, of all its centuries-old deposits, of all the treasured richness of its dialects. Therefore, the dialectal activity, which continues always, is no longer a danger: it can no longer tear the language apart. Now, instead of killing it, the dialects invigorate it, just as the breezes from the fields invigorate the blood of the cities’

“Crear una llengua literària és l’obra cabdal d’un poble [...]. Quan del conjunt de dialectes d’una llengua en sobresurt un i s’enlaira sobre els altres, i dóna forma a les creacions literàries i filosòfiques i jurídiques de la raça, aquella llengua en estat natural de què parlàvem, aquella incoherent munió de parlars dialectals, queda articulada, és convertida en una unitat vivent, en un centre de relligament i coordinació i depuració de totes les forces lingüístiques de la raça, de tots els seus dipòsits seculars, de totes les riqueses atresorades en els seus dialectes. Llavors, l’activitat dialectal, que sempre continua, deixa d’ésser un perill: ja no pot portar a la tritració de la llengua. Ja, en comptes de matar-la, la tonifica, com els aires del camp tonifiquen la sang de les ciutats” (Prat de la Riba 2000, 697).

In 1914, the same Prat de la Riba was elected president of the Mancomunitat de Catalunya, the first assembly since the abolition of Catalan institutions in the early 18th century to have some degree of political control over the four ‘provincies’ of Catalonia. The recent IEC spelling norms would be used as standard in the new

administration. That same year, the columnist Antoni Rovira i Virgili admitted that the Catalan language was 'still in a period of elaboration, preparation, re-working' ("encara en un període d'elaboració, de formació, de reforja"). But one thing was clear: 'the old literary language has fallen' ("el vell llenguatge literari ha caigut"). The corrupted language of the streets had been corrected with contributions from old Catalan, with words from various dialects, by comparing and contrasting with other Romance languages and by assimilating new words. Thus it was that 'today's Catalan is not a perfect Catalan, but it is much better. This Catalan is alive, fresh, rejuvenated. In fifteen years, much of the road has been travelled' ("el català d'avui és un català no pas perfecte, però sí molt millorat. És un català viu, fresc, rejuenit. En quinze anys s'ha fet un bon tros de camí"). What was needed, Rovira said, were some 'serious' studies into lexicography and dialectology to complete the journey and leave behind the state of confusion once and for all (Rovira i Virgili 1914a, XII–XIII).

The subsequent surge in publications on Catalan linguistics was extraordinary. Still in 1914, the Secció Filològica began publishing the *Butlletí de Dialectologia Catalana*. A work by Pere Pujol, *Documents en vulgar per a l'estudi de la llengua* (segles XI, XII & XIII) (1913) was published by the IEC's newly-formed *Biblioteca Filològica*. Students who had been abroad on scholarships granted by the Barcelona Provincial Council had their doctoral reports published, including Pere Barnils, *Die Mundart von Alacant. Beitrag zur Kenntnis des Valencianischen* (1913); Antoni Griera, *La frontera catalano-aragonesa* (1914) and Josep M. d'Arteaga i Pereira, *Textes catalans avec leur transcription phonétique* (1915). The task of modernizing and purifying was soon facilitated with the *Diccionari ortogràfic* (1917), published by the IEC and drawn up under the editorship of Fabra. And the following year came Fabra's *Gramàtica de la llengua catalana* (1918), which became accepted as standard.

The vast majority of writers began to accept the new rules and adjust their creative practices to what was set out in the texts issued by the IEC and adopted by the renewed administrations, beginning with the Mancomunitat de Catalunya. In 1915, Fabra warned that there was only one right way, which did not admit alternatives: 'The writer who disparages or ignores the work of the grammarian is a bad writer and a bad patriot' ("L'escriptor que menysprea o ignora la tasca del gramàtic és un mal escriptor i un mal patriota", Lamuela/Murgades 1984, 183). Over and above disparity, there had to be a fair and balanced discipline. Thanks to the 'patriotic' zeal of the reviewers trained in the Fabra camp – in particular, Emili Guanyavents and Emili Vallès, working in the Oficina de Revisió d'Originals – from now on the implementation of the system of spelling and grammar laid down by Fabra would go on creating 'the awareness of a language that was unique, monolithic and, regardless of levels or register or geographical variants, the literary language, in which any phenomenon was, without nuance, either right or wrong' ("la consciència d'una llengua única, monolítica, sense distinció de nivells o registres o varietats geogràfiques, la literària, en la qual un fenomen era, sense matisos, correcte o incorrecte", Rico/Solà 1995, 55).

Fabra's deliberations found an excellent platform in the progressive Catalanization of the republican newspaper, *La Publicidad*, a process which was completed in 1922 when it was re-named *La Publicitat* and became openly Catalanist and progressive. In 1919, Fabra began to contribute to the newspaper with a series entitled *Converses filològiques*. Now accepted by public opinion as a bona fide authority on the language, he wrote more than eight hundred articles between 1919 and 1928 (collected in Fabra 2010), producing 'the most important corpus of doctrine on modern Catalan' ("el corpus més important de doctrina sobre el català modern", Segarra 1991, 107). With a style the public found easy to understand, Fabra dispensed his wisdom on the appropriateness or otherwise of words, grammar and pronunciation. He was able to base his conclusions on historical documentation (provided largely by the *Diccionari Aguiló*), on medieval literature (the collection *Els Nostres Clàssics* since 1924), on diatopic criteria (Griera began publication of the *Atlas lingüístic de Catalunya* in 1923) and on current use.

In 1927, Emili Vallès, a direct disciple of Fabra, began publication, in instalments, of the *Diccionari català-castellà-francès Pal·las*. In 1930, he went on to produce a *Diccionari de barbarismes del català modern*, which had a decisive impact on subsequent selections. 1930 also saw the first of the four volumes of the *Diccionari enciclopèdic de la llengua catalana amb la correspondència castellana* (1930–1935), an adaptation of the "pre-standardization" edition which appeared in 1910. This was the first encyclopedia in Catalan which, in both form and substance, adopted the assumptions inherent in the understanding of Catalan as a "llengua nacional".

In the end, the culmination of Fabra's work to standardize the language came in 1932 with the publication of the *Diccionari general de la llengua catalana*. The 51,791 articles that make up this great work were not intended to contain all the words inventoried by the linguist and his collaborators in the years leading up to its publication, but they did impose operational restrictions on the standardized Catalan corpus, expressly excluding archaic words that had fallen into disuse or insignificant dialectal words and Castilianisms. This dictionary would become known as "El Fabra" and was to set the standard for Catalan until the last few years of the 20th century.

Using the development of literary Catalan as stipulated by the IEC as their benchmark, the main regional varieties of the language underwent a process of considerable formal convergence. In French administered Catalonia, Pau Berga published a book of poems in 1913 called *Mare-Terra* using Catalan that was practically unlocatable. At the time, Raymond de Lacvivier, notary of Elna, published *Les Règles orthographiques de la Langue Catalane*, following the IEC. In 1915, Mossèn Esteve Casespnce brought to Barcelona the *Rondalles* of the North-Catalan region of Vallespir, which were perfectly regulated, and added precise vocabulary correspondances. In 1919, the teacher Lluís Pastre demanded that the Perpinyà publishers of the *Revue Catalane* respect the IEC norms (Bonet 2016, 234). From 1921 onwards, the *Colla del Rosselló*, a group of writers from that area, proposed some essentially conservative 'Rossellonese rules' ("normes rosselloneses") that would eventually turn into unifying solutions,

thanks especially to the intervention of the poet and linguist Carles Grandó. Also beginning in 1921, the poet Josep Sebastià Pons, also from Rosselló, began to publish his work in Barcelona, where he was recognized not only for the indisputable quality of his work, but also for adapting his slightly different variety of Catalan to the general norms. Ultimately, in 1936, the intellectual and nationalist Catalan-Rosselló youth movement founded the magazine *Nostra Terra*, with the unconditional acceptance of Catalan as a 'national' interstate language.

Meanwhile in Valencia, the tendency, entrenched for centuries, towards a use of the language uncoupled from generic Catalan, would begin to fade as a result of the progress of the IEC and its regulations. In 1915, the weekly magazine *Patria Nova* adopted the 1913 rules. In 1916, a *Biblioteca Valenciana* was launched, the first volume being the novel *De l'horta i de la montanya*, by Eduard López-Chavarri. Although the prose was lexically and morphologically distinctive, it was still clearly in consonance with the classical language. Ernest Martínez Ferrando, meanwhile, published the collection *Les llunyanies suggestives i altres proses* (1918), sticking rigorously to Fabra's edicts. Subsequently, the teacher Bernat Ortín Benedito published his *Gramàtica valenciana* (1918), with the aim of bringing the essential parts of the Catalan spelling agreement to primary education. Aware of all those developments, Fabra invited writers from Valencia to undertake the work of purification in parallel to the work going on in Catalonia, which would result in the recovery of the old literary unity of the language (Fabra 1980, 147–148).

In the 1920s, young sympathizers of the first wave of Valencian cultural nationalism, with one eye firmly on Barcelona, endorsed the decidedly unitary trend. Around 1927, Miquel Duran i Tortajada published a couple of regionalist novels by Vicente Blasco Ibáñez translated into a Valencian that fully adhered to the 1913 spelling rules. The weekly instalments of *Nostra Novela* (1930–1931) put into practice the decision tacitly agreed within Valencian literary circles and succeeded in popularizing it. In 1930, the Valencian teacher Lluís Revest i Corzo published *La llengua valenciana. Notes per al seu estudi i conreu* in which – avoiding the age-old question of the name of the language itself – he declared 'without hesitation, the indissoluble unity of the language of Valencia, the Balearic Islands and Catalonia' ("sens vacil·lacions la unitat indissoluble de la llengua de València, Balears i Catalunya", Revest i Corzo 1930, 15).

In 1932, with the vast majority of the country's luminaries having signed the agreement known as the *Normes de Castelló*, in Castelló de la Plana, the period of vacillation was over, and it was unanimously decided to transplant the IEC's writing model to Valencia as a whole. The Alacant-based satirical weekly magazine *El Tio Cuc* (1932–1934) incorporated, under the supervision of Enric Valor, spelling lessons, rules for good pronunciation and, most importantly, a writing style based on the *Normes de Castelló*. In the Valencian capital, the weekly newspaper *El Camí* (1932–1934) or the quarterly *La República de les Lletres* (1934–1936) acted similarly in the context of cultural nationalism. Josep Giner (pseudonym of Guillem Renart i Ferrís) presented *La*

conjugació dels verbs en valencià (1933), but it was the teacher and poet Carles Salvador who emerged as the foremost adapter of Fabra's work in the southern areas of the Catalan-speaking territories, with his *Vocabulari ortogràfic valencià* (1933), *Morfologia valenciana* (1935) or *Qüestions de llenguatge* (1936).

However, the strongest contribution in weakening Valencian idiomatic particularity was provided by a youthful Manuel Sanchis Guarner, a pupil of Ramón Menéndez Pidal, in *La llengua dels valencians* (1933) – the first essay to adapt the verdict of the Romance philologists to Valencian. Despite its brevity, Sanchis Guarner defined the generic personality of an entire territory. It was true, of course, that Valencians did not usually talk about their language as 'Catalan' and had not done so for centuries. But whatever it was, for Sanchis, it was absolutely incontestable that it constituted a genealogical extension of Old Catalan, manifest at that time in an area that went far beyond that of the 'Valencian language', stretching from Salses to Guardamar and from Fraga to l'Alguer.

The history of the adoption of the IEC standards in the Balearic and Pityusic Islands is closely linked to the rise of writers from these islands in the Catalan linguistic-literary scene of the early 20th century. That is why its integration into the general mold did not present any doubts or dissent. In the beginning, there was the influence in Catalonia of the writers of the so-called 'Majorcan School' ("Escola Mallorquina") and their followers: Miquel Costa i Llobera, Joan Alcover, Miquel dels Sants Oliver, Maria Antònia Salvà, etc. The collaboration of islanders (Gabriel Alomar, Joan Estelrich, Miquel Ferrà, Llorenç Riber, etc.) in the cultural work of the Mancomunitat would prove important. Also, between 1928 and 1936, the magazine *La Nostra Terra* promoted a linguistic Catalanism without dissent.

It seems that in the islands, the standard rules were not problematic, because there were, in fact, no problems to find. In 1931, Francesc de B. Moll, a close collaborator of Alcover, published an *Ortografia mallorquina segons les normes de l'Institut, adequada al llenguatge de totes les Balears*. That same year the Arabist Jaume Busquets produced a *Curs pràctic d'ortografia i elements de gramàtica*. In the throes of the civil war, Francesc de B. Moll even managed to publish in Majorca, which had already fallen to the fascist rebels, his *Rudiments de gramàtica preceptiva per a ús dels escriptors baleàrics*, a testament to the unity of the Catalan language and the completely integrated role played by all the varieties on the islands.

5 The resistance

In 1918, when the standardization of the reformed Catalan appeared to be unstoppable, Fabra warned that there were two dangers that could hinder it. First, there were those who endeavored to write in an archaic style, and second, there were those who wanted the written language to reflect all the dialectal variants currently spoken (Fabra 1980, 143–145). The counterattack against the IEC's conclusions was immedi-

ate and took either one, or both, of these two routes. Just a few days after the *Normes ortogràfiques* of 1913 were published, Jaume Collell, a canon of Vic, unleashed a series of articles in the *Gazeta Montanyesa*, which encapsulated the old spirit of conservatism. For Collell, ‘the leftists who were the first to rebel against the traditional and rational spelling system [...] now want everyone to bow down before their impositions’ (‘els esquerrans que foren els primers de rebel·lar-se contra el sistema ortogràfic tradicional i racional [...] ara volen que tothom s’ajupi a les seves imposicions’, Requesens i Piqué 1994, 119). As well as going against tradition, the rules would reflect ‘the influence of the metropolis’ – a metropolis that, because of its increasing social bilingualism, was far from exemplary (Requesens i Piqué 1994, 126).

Collell was not alone in his criticism of the renovation and the renovators. In 1915, a genuine troop of former trailblazers, led by the promoter of the *Jocs Florals*, Francesc Matheu, formed their own Acadèmia de la Llengua Catalana. Alcover, president of the IEC’s Secció Filològica, considered the responses of such troublemakers to be ‘empty and utter nonsense’ (‘buides i sense cap ni peus’, 1913, 400). In 1916, however, those opposed to the IEC developments published their own version of Catalan spelling rules, *Regles ortogràfiques*, under the auspices of the veteran Reial Acadèmia de Bones Lletres de Barcelona, an institution that would, in fact, become the nucleus of opposition to the IEC. The list of dissidents was neither short nor insignificant. They included unquestionable figures of Catalan letters of the last third of the 19th century who were still active, such as Josep Pin i Soler, Narcís Oller, Apeles Mestres (who recoiled at the idea of changing his name to the “modern” Apel·les) or Víctor Català. Certainly, these were figures from another era, ‘relics of the *Renaixença*’ (Segarra 1985, 350), but their disagreement, apart from stemming from a prevailing personal or aesthetic sense of distaste, was supported by the work of three activists from the more conservative philological circles.

First of all, there was Antoni Bulbena i Tusell, curator of ancient Catalan texts, author of *Lliçons de gramàtica catalana* (1898) – which challenged the system of *L’Avenç* – and the *Diccionari català-francès-castellà* (1905). Scarcely moderating his devotion to the past, Bulbena would publish a *Diccionario catalán-castellano* (1919) and, as a rejoinder, the *Diccionari de les llengües francesa & catalana* (1921). For Bulbena, any Catalan that diverged from that spoken in the 16th century was impure by definition, and therefore abominable. More or less of the same opinion, Anfós Par published the first historical syntax of Catalan in 1923: *Sintaxi catalana segons los escrits en prosa de Bernat Metge*.

However, the most pugnacious of the anti-modernist language theorists was Ramon Miquel i Planas, a conscientious publisher and writer, scrupulous curator of the Catalan, Valencian and Balearic medieval classics in the *Biblioteca Catalana* (1908–1950), as well as in the more popular collection *Històries d’altre temps* (1910–1917). He was the living embodiment of militant anti-Fabrism. From the pages of his highly selective journal *Bibliofilia* (1911–1920), Miquel i Planas attacked anything implying innovation in the Catalan language. His most resounding contribution,

however, came at a conference entitled “Contra la Reforma Lingüística” (May 1918), where, in addition to denouncing the ways in which a ‘journalistic dialect’ (“dialecte periodístic”) was gradually being concocted, he also denounced the spirit underlying all such reform, which he saw as symptomatic of revolution. In 1920, Miquel i Planas published the book entitled *Purgatori del bibliòfil*, which exhibits an unrestrained aversion towards the IEC and Fabra in particular.

The dissidents had various outlets for their thoughts. Between 1913 and 1925, Francesc Matheu re-published all the works of Jacint Verdaguer – a truly popular poet in Catalonia – applying his own orthographic criteria. Matheu also published fascicles of a collection entitled *Biblioteca d'Autors Catalans* (1913–1921), a literary canon which served as an alternative to the then-current *Noucentisme* movement and to the language formalities being established. To sustain the anti-Fabra brigade, Matheu founded the weekly magazine *Catalana* (1918–1926). Another important blow from the ‘anti-regulatory’ (“antinormista”) cause arrived with the publication, in six volumes, of the *Geografia general de Catalunya* (1908–1918), edited by Francesc Carreras Candi, a prominent member of the Acadèmia de Bones Lletres de Barcelona, who had his own particular view on the historical origins of the Catalan language. From within the IEC, there was a persistent resistance to the new spelling. Even as late as 1921, Antoni Rubió i Lluch, who had been the first president, published the second large volume of the *Documents per l'Historia de la Cultura Catalana Mig-*eval** still with the old spelling.

Other authors, while sticking to the IEC standards, were critical of their being applied too strictly or, more specifically, of the excessive zeal of those interpreting the rules. The priest Ricard Aragó (pseud. Ivon l'Escop), promoter of a highly active anti-blasphemy association, championed the spontaneity of Maragall: ‘If the literary language [...] does not adjust itself to the biological laws of popular speech and to the local dialects, adapting to each time, to the product of the variations and changes and shifts of all places, then, that literary language, rather than being a remedy in the standardization and unification will leave it in danger of being impoverished and denatured’ (“Si la llengua literària [...] no s’ajusta a les lleis biològiques de la paraula popular, als dialectes locals, atemperant-se a cada temps a la suma de les variants i canvis i mudances de tots els llocs, aquella llengua literària, en lloc de ser un remei en la generalització i unificació, portarà un perill: l’empobriment i la desnaturalització”, L’Escop 1921, 94). Another cleric, the Jesuit Josep Calveras, went as far as to write a treatise on *La reconstrucció del llenguatge literari català* (1925), which proposed the contiguous nature of regulations and the living language, and condemned the criteria of both the classicalist and the extreme differentialists. Like the traditionalists, Calveras felt that Fabra had ‘turned Catalan sociolinguistic reality into something abstract and now dared to promote a reformed language unnecessarily differentiated from Spanish’ (“feia abstracció de la realitat sociolingüística catalana i per això s’atrevia a promulgar una llengua reformada innecessàriament diferenciada del castellà”, Iglésias 2004, 252).

The IEC's work also provoked rejection from regional quarters. In Valencia, the work of the Franciscan monk Lluís Fullana stands out in particular. Initially involved in the work of Mossèn Alcover, in 1915 Fullana published his *Gramàtica elemental de la llengua valenciana*, at odds with the recent orthographic verdicts from the IEC. Despite considering that Valencian was essentially no different from the language of Catalonia and the Balearics, in his later work Fullana steered towards a differentialism that would become increasingly pronounced. Fullana may not have created the myth of Mozarabism, which claimed that Valencian was a Romance language that derived directly from Latin without any Catalan intervention, but he was its chief promoter. The development of this theory over time is reflected in his works: *Vocabulari ortogràfic valencià-castellà* (1921), *Compèndi de gramàtica valenciana* (1922) and in *Ortografia valenciana* (1932).

In French Catalonia, a similar disaffection was apparent in more traditionalist circles. The IEC rules began to spread at the outbreak of the World War II. Just a few days after the declaration of hostilities, Juli Delpont – who had participated in the 1906 Congress – denounced the supposed conspiracy between official Catalanism in Barcelona and the German Romanists with whom they had lived and sympathized at the beginning of the century. Hence the oft-cited association made by Delpont between orthographic reformism and “la philologie philoboche” in the magazine *Montanyes Regalades* (Verdaguer 2002). From Perpinyà, Josep Calmette, who had also answered the call of Alcover in 1906, wondered: “Pourquoi, sous prétexte d'unification, voudrait-on nous imposer à nous, Roussillonnais, les déformations du véritable catalan qu'a subies le dialecte barcelonais, du fait de l'influence castillane?” (Calmette 1916).

Alcover himself is a special case. Appointed president of the Secció Filològica in 1911, Alcover would progressively reduce his participation in the institution. But it was not until 1918 that Alcover broke all ties with the IEC, taking with him the documentation that he had collected there, with an eye to compiling a comprehensive dictionary. In the article entitled “Què ha d'esser el català literari?” published around the time of the dissemination of Fabra's prescriptive grammar, Alcover attacked those who ‘support the disproved and outrageous thesis that the Catalan in Barcelona should be the standard and law and sign of good Catalan’ (“sostenen la desbaratada i sublevadora tesi de que el *català de Barcelona* ha d'esser norma i la llei i la cifra del bon català, lo català literari”, Alcover 1918, 50). From this point on, Alcover, having returned to Majorca, devoted himself to this battle. One result was a change in the name of the dictionary he was still preparing: it would now be known as *Diccionari català-valencià-balear* (DCVB). The Spanish monarchy would provide funds to carry out the work, with the staff in Barcelona excluded.

With the co-operation of the young Francesc de B. Moll from Ciutadella, the first volume of the DCVB was published in 1930 and consisted of more than 15,000 articles written according to the pre-1913 system. Since it was intended to be a complete dictionary of the language, for entries that were orthographically controversial, Alcover included the note “(IEC)”, with the solution that was generally accepted but

which he himself dismissed. In 1932, when the second volume was almost completed, Antoni M. Alcover died, clearing the way for Francesc de B. Moll to redirect the *DCVB* on the path to unification with the solutions provided by the IEC.

6 Dissemination of “the national language”

The same year the Mancomunitat de Catalunya was launched, Antoni Rovira i Virgili was reviewing the state of the social use of Catalan in Catalonia. Despite the progress made since the institutionalization of political Catalanism, Rovira took note of the diglossia still prevailing in the country. The outward appearance of cities and towns in Catalonia was *Castilian*. In primary schools, the historical language of the people was still largely ignored. In universities, it existed on the margins. In the press, in books or at the theatre, the official language, Castilian, dominated the market. Rovira concluded by saying that Catalonia was ‘among all of Europe’s dominated nationalities, the one whose language rights – the soul of the country – were most trampled on’ (“una de les nacionalitats dominades d’Europa on els drets de la llengua –ànima de la pàtria– estan més trepitjats”, Rovira i Virgili 1914b, 21).

Shortly after these words were printed a somewhat inflamed senator declared “Cataluña debe seguir el ejemplo de sus dos entrañables compañeras en el curso de nuestra historia: Valencia y Aragón, que se han incorporado por completo a la vida nacional española” (LlenCat 1916, 60). In the ensuing protests, there was a heated debate in the Barcelona City Council, with a view to restoring Catalan as the language of the administration. This new atmosphere also precipitated the founding of a new association, *Nostra Parla*, whose aim was to promote Pan-Catalan ideals with publications, celebrations and maps and which was credited with popularizing the sense of linguistic unity more or less throughout the territory (Graña i Zapata 1995).

The social penetration of the reforms emanating from the Secció Filològica of the IEC was conducted on several fronts. The most visible of all was literature. In 1917, sectors close to *La Veu de Catalunya* founded the Editorial Catalana with the aim of ‘nationalizing’ the recently standardized language. All manner of high-quality publications were issued, such as the magazines *D’Ací i d’Allà*, *Agricultura*, *Economia i Finances*, the booklets of the *Minerva* collection, which were widely circulated, or the *Enciclopèdia Catalana*, comprising short monographs. They also promoted the publication of a *Biblioteca Literària*, with translations of foreign authors and a *Biblioteca Catalana* with works by local authors. With large print runs and good distribution networks, *Editorial Catalana* was the most prolific company publishing in Catalan that had ever existed. The problem that this new company immediately faced was with the recovery of writers who had been published before the Fabra reform. As one of them put it, ‘given the general acceptance that the rules have achieved, the writer who does not adopt them is condemned to have no readers, since future generations, accustomed through the press to the new orthography, shall reject any books in the future

that are written in the old style' ("donada l'acceptació general que les normes promulgades han assolit, l'escriptor que no les adopti es condemna a no tenir lectors, car les noves generacions, avesades per la premsa a la nova ortografia, rebutjaran en l'esdevenidor els llibres que trobin grafiats a l'antiga", Maseras 1996, 88).

The *Biblioteca Catalana* (1919–1922) launched into a systematic revision of modern Catalan classics (Narcís Oller, Santiago Rusiñol, Joaquim Ruyra, Prudenci Bertrana) which, in this way, would become definitive versions, not only with a different spelling from the original, but often with a highly modified morphosyntax and lexicon (Rafanell 2017). The great novel *La vida i la mort d'en Jordi Fragonal* (1912) was written by Josep Pous i Pagès, who had found it very hard to accept the *Normes ortogràfiques* but amended the work himself from start to finish in 1926.

Another area of socialization of the IEC standards involved children's publications. Josep M. Folch i Torres set to work updating the magazine *En Patufet* and related publications (*Biblioteca Patufet*, *Pàgines viscudes*, *Biblioteca Gentil*, etc.) all of which were hugely popular. This had the immediate effect of familiarizing young children with modern Catalan and served as an important counter to the castilianizing effects of the standard school curriculum. It was undoubtedly true that the official education system continued to devalue Catalan and, in response, there was a resurgence of the Associació Protectora de l'Ensenyança Catalana (APEC) which had emerged in the heat of the *fin-de-siècle* *Renaixença* movement. Teacher training colleges began allocating teaching positions specializing in Catalan grammar. For the training of the teachers, there were the magazine *Quaderns d'Estudi* (1915–1924). And for the children, the weekly magazine *La Mainada* (1921–1923), collected texts from the best writers of the time (Joan Salvat-Papasseit, Ignasi Iglésias, Josep Carner, Carles Riba, etc.).

Schools introduced textbooks, such as *Garba. Antologia de les lletres catalanes*, by Lluís G. Pla (1915) expressly adapted to IEC standards, bringing Catalan back into the classrooms. Some books were updated reprints of pre-1913 works: *Aplec. Models en vers i en prosa del nostre renaixement per a ús de les escoles de Catalunya, Mallorca, València i Rosselló*, by Anton Busquets i Punset (1916); *Prosa i vers. Lectures morals i cíviques per a nois i noies*, by Manel Marinel·lo (1917); *El trovador català. Llibre de lectura en vers destinat als col·legis de nois i noies de Barcelona*, by Antoni Bori i Fontestà (undated). In 1918, APEC edited an adapted version of Fabra's recently published standards and also issued all manner of pedagogical manuals: on Catalan history (Ferran Soldevila and Ferran Valls i Taberner), on the history of modern Catalan literature (Manuel de Montoliu), on arithmetic (Lluís G. Castellà), accounting (Ferran Boter) and the geography of Catalonia (Pere Blasi). For younger children, Pau Romeva published *Sil·labari català* (1922), bringing Fabra's own 1904 version up to date.

The proper use of modern Catalan was not limited to writing. Fabra had already set out various guidelines for good diction scattered throughout his work and in 1929 these were gathered together and summarized in *Exercicis de llengua catalana*, by Jeroni Marvà (pseud. Artur Martorell and Emili Vallès). The subordination to the official Spanish was an obstacle to the assimilation of the newly-emancipated Catalan

prosody. There were however some areas in which it could be executed. Refined oral Catalan could be heard in the theatre, especially with the collusion of the younger generation of playwrights in favor of the reform, including the moralistic Folch i Torres and the worldly Josep M. de Sagarra. There was, moreover, sacred oratory, in the form of the catechesis, prayers and sermons. From the first Congrés Litúrgic de Montserrat there emerged *Eucologi* (1915), followed by a people's missal, *Missal del poble* (1919). The ecclesiastical publishing house, *Foment de Pietat Catalana*, brought out a collection of formularies (*El llibre de la comunió diària* [1919] or the *Missal romà per a ús dels fidels de les diòcesis de Catalunya* [1926]) which reinforced both written and pronunciation standards, learned in the seminaries and promulgated from the pulpits.

In the early 1920s, Catalonia was engaged in a determined process of recovery from the uneven diglossia. But then, in September 1923, the Captain General of Barcelona, Miguel Primo de Rivera, carried out a *coup d'état* with the acquiescence of King Alfonso XIII. Thus began a dictatorship as severe as it was unstable, although there was one over-riding element underlining the intentions of the new regime: a marked anti-Catalanism, particularly with regard to its linguistic and cultural elements. In Catalonia, the dictatorship devoted its full attention to preventing or limiting the use of Catalan at all levels of communication where this was possible. Shortly after the military coup, town halls and councils were relieved of their responsibilities. Fabra lost his position at the Mancomunitat and was replaced by Alcover before the *Mancomunitat* itself was dissolved. In 1925, the Institut d'Estudis Catalans was renamed the *Instituto de Estudios Superiores de Cataluña*.

Despite the immediate uproar, Catalan would be banished from the places of importance that it had only recently re-conquered. The theories opposing the 1913 rules found fertile ground thanks to newspapers sympathetic to the new regime publishing all kinds of legitimizing arguments and the language standardized by the IEC was described as “un extraño dialecto catalanista” (Roig Rosich 1992, 197). In 1926, Primo de Rivera laid out his reform of the *Academia Española* in order to assimilate the ‘regional languages’. It was a purely compensatory exercise serving to secure the co-operation of several figures who had been more or less left out of the politics of the Mancomunitat. While Fabra refused to participate, Eugeni d'Ors (erstwhile Catalanist, now transplanted to Madrid), the Majorcan priest Alcover and the Valencian ‘Mozarabist’ Lluís Fullana agreed to comply.

Nevertheless, during the six long years of Primo de Rivera's dictatorship, the orthographical, grammatical and lexical rules that had become standard were, in fact, fortified. Francesc Cambó even wrote that the Catalan language ‘had never been so strong’ (“no havia estat mai tan forta”) as it was during the dictatorship (Cambó 1930, 30). Despite the censorship, the various publications distributed in standard Catalan, which had been created and produced in Catalonia in the previous decade, went from strength to strength, bolstering a literary circuit and a cultural market. Thanks to the patronage of Cambó, in 1923 there appears the collection of Greco-Latin classics *Bernat Metge*, homologous of the French *Collection Budé* and American *Loeb Classical*

Library. Those newspapers existing before the coup, such as *La Veu de Catalunya* or *La Publicitat*, were joined by new ones, such as *La Nau* (1927–1929), *L'Opinió* (1928) or *El Matí* (1929). Longstanding periodicals, such as *La Revista* or *D'Ací i d'Allà*, were joined by *Revista de Catalunya* (1924), *La Paraula Cristiana* (1925), *La Dona Catalana* (1925), *La Nova Revista* (1927–1929), *Mirador* (1929), *Imatges* (1930) and more. From sports to fashion, radio broadcasting to the decorative arts, all fields of modern communication would be represented in print in a neutral and recognizable Catalan. This was the time when the weekly magazines, *L'Esquella de la Torratxa*, *La Campana de Gràcia*, the dirty *Papitu* and the sporting *Xut!* (1922–1936) fused vulgar discussion with more selective language, a sure sign of the extent of the broadening of the IEC project. In the field of formal divulgation, the Editorial Barcino, by then a new publishing label, launched the “Col·lecció Popular Barcino” in 1925 (ca. 160 volumes up to 1938) fifteen years before the first issue of the French collection “Que sais-je?”.

Thus, the period of Primo de Rivera's anti-Catalan government led, paradoxically, to the irreversible triumph of Catalan being understood as “llengua nacional”. In Catalonia, at least, now was the time when, thanks to the Catalan language press, the idea of linguistic unity and virtuous unification was promoted every day.

‘Right now, that is, in the times we live in, a phenomenon is taking place in Catalonia which many are not aware of [...]. The phenomenon in question is linguistic in nature. The Catalan language, which has many dialects and many varieties has, since some time ago, within Catalonia proper, entered into a period of relatively rapid and highly resolute unification’.

“Ara mateix, és a dir, dins el temps en què vivim, es realitza a Catalunya un fenomen del qual molts no s'adonen [...]. El fenomen a què ens referim és de caràcter lingüístic. La llengua catalana, que té nombrosos dialectes i múltiples varietats, ha entrat d'algun temps ençà, dins la Catalunya estricta, en un període d'unificació decidida i relativament ràpida” (*LP*, 19-III-1925).

The enthusiasm of Catalanist intellectuals led Fabra to fear that the example of the process of refining Catalan, imitated by like-minded peers in the north could end up converting the Catalan language into ‘one variant more of a re-discovered great Occitan language’ (“una variant més de la gran llengua occitana retrobada”, Lamuela/Murgades 1984, 194). However, if only fifteen years previously, the dominant impression had been that the regulations were provisional, ‘in recent years the progress of Fabra's work on grammar has been so great that today it appears as definitively victorious’ (“en els últims anys l'avenç de l'obra gramatical de Fabra ha estat tan gran, que avui apareix ja com a definitivament victoriós”, *LP*, 2-IX-25).

General Primo de Rivera's regime abolished the standardizing institutions but not the implementation of their work. In 1929, shortly before entering an economic crisis, there were twelve newspapers published in Catalan in Catalonia. *Ràdio Associació de Catalunya* was on the verge of beginning to broadcast exclusively in Catalan. Catalan literature was enjoying great success with the launch of the publisher Proa – suffused with a genuine cosmopolitanism – or the *Premi Crexells* for fiction. Despite a wave of migration from the south of Spain, which only increased as preparations went on for the Universal Exposition of 1929, Barcelona was generally living in her own language.

Proof of this instinctive normality, immune to government procedures, is provided by the French observer Léon Rodin, who declared that Catalan “est aujourd’hui un véhicule de culture universelle et cela c’est produit précisément au moment où il est interdit de l’employer dans les actes officiels” (*Le Temps*, 5-VI-1929).

With the republican revolution in April 1931, the Catalan language obtained the legal recognition that it had been lacking for more than two centuries. The Statute of Autonomy voted for by the Catalan people established it as the only official language for internal purposes, although the text finally approved by the Spanish Parliament would see it share official status with Spanish. The repercussions of this legislative change would, in any case, be immediate and unmistakable. There was a sudden transformation of public road signs. The names of towns and cities, streets and squares and shop signs were switched to Catalan. The use of Catalan in primary education and, to a lesser extent, in secondary schools, led to a surge in the publication of school books in standardized Catalan. In 1933, the creation of the *Universitat Autònoma de Barcelona* opened up a path for Catalan which had, until then, been closed. Catalan could now be used in the administration of justice; court cases, tribunals and hearings in Catalonia could be carried out in Catalan, and the language was also to be used by notaries, who, since the mid-19th century had worked exclusively in Spanish. Hence the publication of manuals of administrative language, compiled by the writer Cèsar August Jordana. The Valencian doctor Manuel Corachan did the same with the *Diccionari de medicina* (1936). As a result, by 1933 there were 27 newspapers being published in correct and corrected Catalan. There were also two radio stations broadcasting in the language of the country. In one of them, Artur Balot, a teacher appointed by the autonomous government, disseminated the work of the IEC and recreated the ideal elocution conceived by Pompeu Fabra. This whole process of standardization led to the creation of the figure of the Catalan language specialist: translator, reviewer and teacher.

At the beginning of May 1934, a manifesto entitled “Desviacions en els conceptes de llengua i de Pàtria” was published, signed by practically all the leading Catalanist intellectuals, headed by Pompeu Fabra. Despite being rather lengthy, it is the most widely circulated text on linguistics in Catalan history, as it would be printed by the vast majority of Catalan publications of the time. The signatories denounced the two ‘errors’ that in their view might harm the standardization of Catalan:

‘first, the conception of our homeland as consisting solely of the territory governed by the current Generalitat [The Catalan Autonomous Government], that is, reducing it to one of the regions it comprises, as a result of a weakening of national consciousness; second, the conception that confers our homeland a greater area than it has, and which it has never had, due to a confusion of Catalanism with Occitanism’;

“primer, la concepció de la nostra Pàtria com a formada únicament pel territori de l’actual Generalitat, és a dir, la seva reducció a una de les regions que la integren, fruit d’un afebliment de consciència nacional; segon, la concepció que dóna a la nostra Pàtria una extensió excessiva, que no ha tingut mai, provinent d’una confusió de catalanisme amb occitanisme” (Rafanell 2006, 1141).

In other words, on the one hand, they denounced the restriction of the Catalan language to just Catalonia and, on the other, its dilution within the Occitan language area. In any case, the manifesto served to remind the cultured people of the country, at a time when they thought the language was in the ascendancy, of the unity and uniqueness of their language.

A few weeks after the release of this manifesto, Pompeu Fabra presided over the *Jocs Florals de Barcelona*, spelt *Jocs* rather than *Jochs* for the first time in modern history. In his speech, Fabra reviewed the history of linguistic reform, its principles and the obstacles which were overcome in its application. He also recalled the ‘very understandable’ (“ben explicable”) opposition to such rules. Now, however, the path he had etched out more than four decades had been cleared. Although the work of refinement ‘is still unfinished’ (“està encara inacabada”), and its dissemination ‘was just beginning’ (“tot just comença”), Fabra concluded his speech in front of the most conservative audience of Catalanism by saying ‘today we can be glad to have achieved not only this orthographic unity, but also that linguistic unity which has made it possible to elevate Catalan to the official language of Catalonia’ (“avui podem estar joiosos d’haver aconseguit, no solament aquesta unitat ortogràfica, sinó encara aquella unitat lingüística que ha fet possible l’elevació del català a llengua oficial de Catalunya”, Fabra 1934, 26–27). Even in 1938, with Franco’s troops occupying Catalonia, a shortened version of the *Diccionari enciclopèdic català* was published. A few months later, however, with the “national” army now occupying the whole of Catalonia, the nationalization of the Catalan language was once again – and this time for many years – cut short.

7 Bibliography

- Alcover, Antoni M. (1901), *Diccionari de la llengua catalana. Lletres de convit que a tots els amichs d’aquesta llengua envia mossen Antoni M.^a Alcover, Pre. Vicari general de Mallorca*, Mallorca, Estampa de Felip Guasp.
- Alcover, Antoni M. (1913), *Els inemics de les Normes Ortogràfiques van de rotatuda*, Bolletí del Diccionari de la Llengua Catalana 7, 398–400.
- Alcover, Antoni M. (1918), *Què ha d’esser el català literari?*, Bolletí del Diccionari de la Llengua Catalana 10, 49–54.
- Balcells, Albert/Pujol, Enric (2002), *Història de l’Institut d’Estudis Catalans*, vol. 1: 1907–1942, Barcelona, Institut d’Estudis Catalans.
- Bonet, Lluç (2016), *Lluís Pastre (1863–1927) i el conflicte lingüístic arran de la Gran Guerra*, in: Teresa Dalmau/Marie Grau (edd.), *La guerre des écrivains roussillonnais/La guerra dels escriptors rossellonesos*, Perpignan, Presses Universitaires de Perpignan, 207–243.
- Calmette, Joseph (1916), *La langue catalane du Roussillon*, Montanyes Regalades 2, 41–42.
- Cambó, Francesc (1930), *Per la concòrdia*, Barcelona, Llibreria Catalònia.
- Canivell, Eudald (1890), *La rutina del català escrit*, L’Avenç 2 (31-VII), 156–191.
- Casas-Carbó, Joaquim (1892), *La reforma lingüística*, L’Avenç 4 (juliol), 222–224.
- CILC (1908), *Primer Congrés Internacional de la Llengua Catalana. Barcelona. Octubre de 1906*, Barcelona, Joaquim Horta, 1908.

- Fabra, Pompeu (1907), *Sobre diferents problemes pendents en l'actual català literari*, Anuari de l'Institut d'Estudis Catalans 1, 352–369.
- Fabra, Pompeu (1934), *Discurs del President*, in: *Jocs Florals de Barcelona. Any LXXVI de llur restauració*, Barcelona, 19–29.
- Fabra, Pompeu (1980), *La llengua catalana i la seva normalització*, Barcelona, Edicions 62/La Caixa.
- Fabra, Pompeu (2005), *Obres completes*, vol. 1: *Gramàtiques de 1891, 1898, 1912*, Barcelona / València / Palma, Proa/Edicions 62/Moll.
- Fabra, Pompeu (2010), *Obres completes*, vol. 7: *Converses filològiques*, Barcelona/València/Palma, Proa/Edicions 62/Moll.
- Fullana, Lluís (1915), *Gramàtica elemental de la llengua valenciana*, Establiment Tipogràfic Domènech, València.
- Graña i Zapata, Isabel (1995), *L'acció pancatalanista i la llengua: Nostra Parla (1916–1924)*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Iglésias, Narcís (2004), *Una revisió de Fabra, una crítica a la norma. L'obra lingüística de Josep Calveras*, Girona, Curbet.
- Labèrnia, Pere (1839–1840), *Diccionari de la llengua catalana: ab la correspondència castellana y llatina*, 3 vol., Barcelona, Estampa dels Hereus de la V. Pla.
- Lamuela, Xavier/Murgades, Josep (1984), *Teoria de la llengua literària segons Fabra*, Barcelona, Quaderns Crema.
- L'Escop, Ivon (1921), *La Paraula en l'Escriptura, en la Gramàtica i en les Acadèmies*, Barcelona, Políglota.
- LlenCat (1916), *La llengua catalana a l'Ajuntament de Barcelona. Debat de 22 de febrer de 1916. Antecedents i conseqüències*, Barcelona, Guinart i Pujolar, impressors.
- Llombart, Constantí (1887), *Ensayo de ortografía lemosino-valenciana*, in: José Escrig y Martínez, *Diccionario valenciano-castellano*, vol. 1, València, Librería de Pascual Aguilar, XV–XXXII.
- Lluch, Ernest (1995), *Producció de llibres en català (1476–1860): la història de la “morta viva”*, L'Avenç 189, 22–27.
- LP = *La Publicitat* (1922–1939).
- LV = *La Vanguardia* (1881–).
- LVC = *La Veu de Catalunya* (1899–1936).
- Maseras, Alfons (1996), *Vida de Narcís Oller*, Tarragona, El Mèdol.
- Matheu, Francesc (1880), *Novas*, La Il·lustració Catalana, 1a època, núm. 16 (10-XII), 126–127.
- Milà y Fontanals, Manuel (1890), *Estudios de lengua catalana*, in: *Obras completas del doctor Manuel Milá y Fontanals*, vol. 3: *Estudios sobre historia, lengua y literatura de Cataluña*, Barcelona, Librería de Álvaro Verdaguer, 510–544.
- Miracle, Josep (1968), *Pompeu Fabra*, Barcelona, Delos-Aymà.
- Ors, Eugeni d' (1908), *Pròleg*, in: Guerau de Liost, *La Montanya d'amethystes*, Barcelona, Octavi Viader.
- Prat de la Riba, Enric (2000), *Obra completa*, vol. 3: 1906–1907, edd. Albert Balcells/Enric Pujol, Barcelona, Institut d'Estudis Catalans/Proa.
- Puiggarí, Pierre (1852), *Grammaire catalane-française à l'usage des français, obligés ou curieux de connaître le catalan, des linguistes et des amateurs de la langue romane*, Perpignan, J.-B. Alzine.
- Rafanell, August (2006), *La il·lusió occitana. La llengua dels catalans, entre Espanya i França*, Barcelona, Quaderns Crema.
- Rafanell, August (2011), *Notícies d'abans d'ahir. Llengua i cultura catalanes al segle XX*, Barcelona, A Contravent.
- Rafanell, August (2017), *Les reedicions de la “Biblioteca Catalana” (1919–1922)*, Estudis Romànics 39, 107–136.
- Requesens i Piqué, Joan (1994), *Jaume Collell i la llengua catalana. Selecció de textos*, Vic, Eumo/Universitat de Girona/Estudis Universitaris de Vic.

- Revest i Corzo, Lluís (1930), *La llengua valenciana. Notes per al seu estudi i conreu*, Castelló de la Plana, Sociedad Castellonense de Cultura.
- Rico, Albert/Solà, Joan (1995), *Gramàtica i lexicografia catalanes: síntesi històrica*, Valencia, Universitat de València.
- Roig Rosich, Josep M. (1992), *La dictadura de Primo de Rivera a Catalunya. Un assaig de repressió cultural*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Rovira i Virgili, Antoni (1914a), *Diccionari català-castellà & castellà-català. Compost en presència dels diccionaris publicats fins ara i de les principals obres literàries i filològiques catalanes, especialment les d'En Pompeu Fabra, i enriquit amb alguns milers de mots que encara no s'havien catalogat*, Barcelona, Antoni López.
- Rovira i Virgili, Antoni (1914b), *La nacionalització de Catalunya*, Barcelona, Societat Catalana d'Edicions.
- Segarra, Mila (1985), *Història de l'ortografia catalana*, Barcelona, Empúries.
- Segarra, Mila (1991), *Pompeu Fabra*, Barcelona, Empúries.
- Verdaguer, Pere (2002), *Història d'una ruptura nord-catalana greu: la revista "Muntanyes Regalades"*, *Revista de Catalunya* 176, 112–130.
- Yxart, José (1890), *El año pasado. Letras y artes en Barcelona*, Barcelona, Librería Española de López.

August Rafanell

14.2 From Pompeu Fabra to the Present Day: Language Change, Hindrance to Corpus and Status Planning

Abstract: The victory of Franco's armies in the Spanish Civil War (1936–1939) led to the prohibition of the use of the Catalan language in Catalonia, Valencia and the Balearic Islands. However, over the course of almost forty years, the Franco dictatorship modulated its repressive policy, with a tolerance that in no case allowed the recognition of Catalan which had been obtained during the period of its nationalization. The corpus of the language had to overcome this political repression as well as the demographic changes experienced by the Catalan-speaking territories in Spain resulting from the massive immigration of allophone populations. Meanwhile, in France the spread of Catalan was interrupted, primarily as a result of the outcome of World War II. The progressive devaluation of the historical language in the Catalan-speaking countries ultimately affected the unity of the language.

Keywords: Francoism, repression, linguistic tolerance, standardization, Catalan language unity

1 Times of silence (1936–1945)

In the Catalan-speaking territories, the outcome of the Spanish Civil War (1936–1939) meant not just an immense human trauma, but also a moral one. After the death and destruction caused by the conflict itself, there now came the retribution wrought by the victors upon the defeated: more deaths, more imprisonments, more persecution. For those Catalan speakers not in exile, the victory of the Francoist forces also led to a kind of civil death. Their language, which had virtually become the only language of the Catalanian administration, was abruptly banished from all areas of life. The key element that the victors of the war sought to destroy was republican legislation which, among other things, had upheld the hard-won status of the Catalan language. In declarations to the *Jornal do Brazil* made in January 1938, General Franco made it clear what kind of national unity he wanted: “la queremos absoluta, con una sola lengua, el castellano, y una sola personalidad, la española” (Benet 1978, 70).

Upon entering Catalan territory in April 1938, the occupying forces abolished the 1932 Statute of Autonomy and the co-official status of Catalan established therein. As the rebel army advanced further into Catalonia, warnings began to proliferate, such as “Si eres español, habla español”, or “Por Dios y por España, habla el idioma del

<https://doi.org/10.1515/9783110450408-022>

Imperio". The *Catecismo patriótico español*, prescribed for use in the Francoist schools, spelled it out clearly: "En España se habla sólo la lengua castellana" (Menéndez-Reigada 2003, 40). Newly arrived in the recently-occupied Barcelona, the university professor Ángel González Palencia, declared: "Ha de mantenerse la unidad lingüística como instrumento de poder, seguros, como Nebrija decía con razón, de que la lengua es compañera del imperio." Anyone blind to this fact needed to be convinced: "es una locura prescindir de una lengua con la que se entienden ochenta millones de hombres en el mundo, para usar otra que no sirve más que para andar por casa" (Benet 1978, 356).

The military dictatorship that emerged from the Spanish Civil War set about eradicating the efforts made over previous years to restore Catalan to the category of a "national language". Hence, in those Catalan-speaking territories that had remained on the margins of such efforts, the linguistic repression was not quite as severe. In Majorca, under Francoist control from the beginning of the conflict, any manifestations that evoked a link with the peninsular Catalan language and identity were suppressed; however, expressions in the local "dialect" were sporadically tolerated. In Valencia, Catalan would even be employed in the celebrations of the triumphant arrival of the Castilian dictatorship (Ballester 1992, 37–43). In all cases, however, the Spanish language was imposed in public life by law – and was imposed with particular severity wherever most action had been taken to reverse the established diglossia before the war. Hence the zealous cries, from the very first day of the occupation of Catalonia, for "re-hispanization", with the aim of banishing the centuries-old language of the Catalans from all forms of social expression.

The work that was now to be completely undone dated back further than the Second Republic and the resulting political autonomy for the Catalans. It had begun much earlier as part of the Catalanism of the late 19th century and, with regard to linguistic-cultural institutionalization, it centered around the efforts of the Institut d'Estudis Catalans (IEC). Once suppressed, the IEC's leading lights either kept silent or went into exile. Among the exiles was Pompeu Fabra, whose work was vilified in the press of the new regime as "aquellos montones de 'ii' latinas y de 'ensems' que [...] tantos ríos de sangre tenían que costar" (Pérez i Vallverdú 2009, 233).

In Catalonia, at the beginning of 1939, the spoken language, the mother tongue of the vast majority of the population of all classes was rendered invisible. Catalan-language newspapers and magazines were closed down. Shop signs and business documentation appeared only in Spanish. The whole administration was transformed, from top to bottom. Place names, which had been standardized by the IEC's Office of Toponymy and adopted by the autonomous government in the 1930s, were replaced by the previous official Spanish versions. Catalan first names ceased to be accepted by the civil registry and became confined strictly for use in the home. In the early days, Spanish was even imposed on religious ceremonies. In the field of education, there was total regression. The stacks of Catalan books which filled classrooms during the 1930s and set out the guidelines of language use disappeared from circula-

tion and were replaced by the spelling books of fascism and National-Catholicism. In the schools of Barcelona, 1,200 of the 1,800 teachers in active employment before 1939 were dismissed or forced into exile. About 700 teachers from outside Catalonia were brought in to take their places. As for universities, almost half the staff were imprisoned, exiled or purged.

During the first stages of the Franco dictatorship, any new publication whatsoever in Catalan was banned (Gallofré i Virgili 1991, 51–82). The only books and pamphlets produced in Catalan were either clandestine or sidestepped the ban with fictitious imprints that claimed the works had been published either in a foreign location or prior to the regime change. Between 1939 and 1944, according to calculations by Albert Manent and Joan Crexell (1988, 12), some 160 titles were published outside the law. The fact that only very few exceptions managed to avoid being outlawed illustrates the full extent of the persecution. However, once the initial repression began to wane, governors in Majorca and Valencia began to make small, indulgent gestures. In 1941, the re-publication of *La ciutat de Mallorques* by Miquel dels Sants Oliver was permitted, albeit with its archaic title. And in 1942, by the same author, came *L'hostal de la bolla*, a story characterized by a certain dialectal and folkloric bent, but nevertheless, the first literary book printed in Catalan in the postwar period. In 1943, an ex-combatant of the Francoist forces, Miquel Dolç, was given permission to publish a book of original poems, *El somni encetat*.

Something similar took place in Valencia. Commonplace printed texts, such as the leaflets prepared for the major fiestas, *Les Falles*, allowed the people to catch glimpses now and then of their own language, even if it was subliterate of dubious quality. The 1943 almanac published by the newspaper *Las Provincias* included some higher quality texts in Catalan. In contrast, such texts would not have been allowed in any shape or form in Catalonia. In the immediate post-war period, Catalan literature could barely eke out an existence on the periphery – where the authorities saw no signs of renewed nationalist sentiment. In Valencian cities, after all, the bourgeoisie had mostly switched to Spanish as the language used in the family home, something which was much less common in Catalonia and far less still in the Balearic Islands.

The enforced silencing of public expression in Catalan was to be used by the Franco regime as a purgative of the preceding years, but the repressive tendencies of Francoism in the early 1940s cannot be fully understood without looking at the contemporary context of the Nazi-fascist axis imposing itself elsewhere. Towards 1943, however, as the balance shifted in World War II, Franco began to curb his more extremist excesses. Soon after, the president of the *Cortes* (parliament) participated in the *Fiesta de las Letras* (designed to replace the abolished *Jocs Florals* of Barcelona) and spoke of the “torpeza insigne pretender extinguir el lenguaje de un pueblo, obra espontánea de su espíritu, y los recuerdos de una literatura como la vuestra, que ya de antaño tiene señalado un destacado puesto de honor en la literatura europea” (LV, 20-VI-1944). Meanwhile, exiled Catalans continued to organize Catalan-only *Jocs Florals* in Buenos Aires, Mexico City, Santiago de Chile, Havana, Bogotá, Montpellier, London, Paris, etc.

When the Nazi-fascist axis collapsed in 1945, the Franco dictatorship was forced to rein in its initial impulses still further. All manner of material printed in Catalan began to slip through the cracks and, in some cases, was openly tolerated. But what kind of Catalan? Was it to be the Catalan standardized by Fabra who, only recently, was considered to be one of the principle agitators against Spanish national unity? Once again, in 1945, the problem arose of choosing a standard model of literary Catalan, just as the country celebrated, with the approval of the regime, the centenary of the birth of the poet and priest, Jacint Verdaguer. There were those who, like the old Catalan philologist Manuel de Montoliu (now transformed into one of the regime's intellectuals) would predict "el retorno al tipo de lengua creada por el poeta, de raigambre popular, y que todos los escritores se han de dar cuenta de que su creación ha de ser función de la creación literaria española y que sólo así podrá ser considerada en todo su valor de variedad regional" (LV, 27-IV-1945).

2 Relative tolerance (1945–1960)

However, for those who had long opposed standardization it was now too late, despite the banishment of Fabra (who was to die in exile in French Catalonia in 1948), which had given them free hand to impose their 'traditional system'. In 1939, Ramon Miquel i Planas was granted permission to publish a book in Catalan, because he was bilingual and extolled the virtues of the victors of the war. The book in question was also written in a style other than that which was regulated by the now banned IEC, whose premises had been raided and whose remnants had fallen under the protection of the Union Académique Internationale. In 1943, permission was granted to publish the complete works of Jacint Verdaguer, a classic among the Catalan classics, provided that the original pre-Fabra spelling was retained and in 1945, Santiago Rusiñol's famous farce, *L'auca del senyor Esteve*, was reprinted under the same criteria.

Nevertheless, the attempt to re-implant a literary Catalan that diverged from the agreed pre-war model would not succeed. 1946 saw the first original book of prose legally published in Catalan in Catalonia: *Mosaic* by Víctor Català, former member of the Acadèmia de la Llengua Catalana. This work was published in perfect Fabrian Catalan. Thus began a second phase in the history of how the Catalan language was viewed by the Franco regime. The prose writer Josep Pla, one of the mainstays of the resurgence of the written language after its abolition, said that 'from the indescribable, unspeakable shipwreck, all that remained standing was one work and one figure: the figure and the work of Pompeu Fabra' ("de l'inenarrable, fabulós naufragi, només queda dempeus una obra i figura: la figura i l'obra de Pompeu Fabra", Rafanell 2011, 423). Despite all of its repressive omnipotence, the Franco regime was unable to liquidate the work of the master.

One of those responsible for safeguarding standardized Catalan in the most difficult of times was Josep M. Cruzet, founder of the Selecta publishing house.

Selecta's low-cost books were soon joined by other companies, such as Millà, Àncora, Dalmau, Aymà, Janés, Barcino, Alpha, Destino or Juventud. In 1946, the Greco-Latin classics of the Fundació Bernat Metge were published once more. The readers who might have been interested in such products had, on the whole, been educated before the war, and had seen for themselves the possibilities of Catalan as a modern and cultured language. In contrast, for younger readers, subject to an educational policy that expressly imposed the Spanish language (confirmed by law in 1945), it was more difficult. The diglossia, which in Catalonia had begun to crumble at the beginning of the century, was once again imposed in the minds of generations who, as time went on, would think more and more of Catalan as a language strictly for use in personal conversations and in the home.

Furthermore, in 1946, theatrical works in Catalan, banned until this point, were authorized in Barcelona. The civil governor of Barcelona wrote that “habían pasado siete años desde que un estado de cosas ya superado había hecho de la lengua catalana un arma más contra España, y parecía ya llegado el momento de incorporarla pacíficamente al acervo común de nuestra riqueza idiomática y dialectal” (Barba 1948, 145). This was the same governor who, in 1947, authorized the use of Catalan in advertisements for the popular fiestas of the Virgin of Montserrat within Catalonia. He was soon removed from office. Meanwhile, in 1951 the City of Barcelona literary awards began accepting original works in Catalan.

Nevertheless, the 1950s were a time of consolidation for the Franco regime due to the international situation at the end of World War II, which placed the Iberian Peninsula in a position of major geostrategic importance. In 1953, the Spanish state signed the first treaties with the United States and established the concordat with the Vatican. In 1955, it was admitted to the UN, the IMF, and the ILO. Consequently, the frontal attack on the language of the Catalans, the Valencians and the Balearic islanders had to be toned down, if only for the sake of Spain's image abroad. In 1950, work was openly resumed in the heart of Barcelona on the *Diccionari català-valencià-balear* (DCVB) by Alcover and Moll – one of the first examples of efforts made by civilians to preserve the language after the Civil War. It was around this time that some in the hierarchy began to speak of the need for a “catalanismo de buen tono”; one that was politically innocuous, rather like the kind practiced in Valencia (Marín i Corbera 1996, 287).

The Franco regime could have copied the moderate language policies that the French government had undertaken in those same years, but it never did. In January 1951, the French Republic established by law (*Loi Deixonne*) that “regional languages” could, optionally, be taught for an hour a week at the end of primary school and in secondary education. The impact of this provision on the Catalan-speaking population in the department of *Pyrénées-Orientales* was minimal. It came at a time when the Catalan language in France was in precarious health; its displacement by French in all areas was already a fact of everyday life. With the Nazi occupation and the war behind them, it appears that the people of Rosselló also wanted to break with the past.

Intergenerational transmission of the historical language came to an end in the cities and towns. Catalan, all of a sudden, was no longer the primary language for most of the Catalans north of the Pyrenees. It was a phenomenon that, although unquestionable, has been overlooked by most observers of the period. But the fact is that in 1955, it was found that the *Loi Deixonne* was not implemented to any significant extent in Rosselló's primary schools. This law – the first of its kind to show tolerance of the local, historical language – would become “la loi du vide” (Berjoan 2011, 238–239). The linguist Albert Dauzat, from Auvergne, was astonished that the new legislation “n’a pas eu de succès auprès des paysans” in France; that is to say, that the only remaining people who retained, to some extent, the “dialect” apparently took no interest in it (Rafanell 2011, 447).

For the Catalans to the south of the Pyrenees, there were no laws permitting the teaching of their own language until the dictatorship ended. At universities, even the scientific study of the language was non-existent. The courses of Romance dialectology taught at the University of Barcelona by the clergyman Antoni Griera – formerly a member of the IEC, who then came into conflict with that institution and was now collaborating with the Franco regime – were no compensation for the ban. The lessons on ancient literature imparted by Martí de Riquer got around the omission of the Catalan tradition by using Occitan literature (reduced to “Provençal”). The *Estudis Universitaris Catalans*, where some of the great teachers who had not been exiled taught Catalan language and literature, had to be carried out in private homes. In these circumstances, at the beginning of April 1953, the VII Congreso Internacional de Lingüística Románica was held in Barcelona, under the auspices of the Regime, with the aim of clarifying the position of Catalan within the Romance family. In fact, it was the second such international congress of Catalan, after the one held in 1906.

The authorities attempted to make the world see that the language policies of the 1950s were not like those of the early 1940s. However, the objective limitations remained very much in place. Catalan was still excluded from schools, radios, newspapers, and any kind of advertising. It would soon also be excluded from the main medium for the transmission of language and values: television. The magazines continued to exist in clandestinity and ended up closing, as in the case of *Poesia* (1944–1945), *Ariel* (1946–1951), *Antologia* (1947–1948), *Dau al Set* (1948–1955) and *Occident* (1950). Moreover, translations of contemporary works into Catalan were prevented, considered unnecessary since they were already in the official language. It was true, however, that in the select world of letters, things were changing. In this decade new, authorized, editorial collections, proliferated: the Biblioteca Raixa, in Majorca, which began with *El que s’ha de saber de la llengua catalana* by Joan Coromines (1954); Sicània, in Valencia, published all kinds of smart little books, with appendices of spelling and lexical clarifications; in Catalonia, there were Albertí and the Club dels Novel·listes, which disseminated contemporary popular literature. Miscellaneous texts appeared to stand in for the suppressed Catalan periodicals: *El Pont*, *L’Ocell de Paper*, *Quart Creixent* and the Valencian magazine *Sicània*. In 1956, Aymà

published the novel *El bon assassí* by Dutchman Antoon Coolen, which was likely the first contemporary literary work translated into Catalan.

The new leaders thought that with these concessions, they had resolved the problem of the language. The official report authorizing the publication of Pompeu Fabra's *Converses filològiques* could not be more definitive in this regard. Having first noted that Fabra “es un catalanista furibundo que ha dedicado toda su vida a la exaltación morfológica y filológica de la lengua catalana”, the censor decided that the book could be published because, “a pesar del matiz catalanista de la obra, como gracias al Movimiento se ha castellanizado, creo que totalmente, la vida catalana y cada día son menos los interesados por estos temas, se considera que no hay inconveniente en que esta obra pueda publicarse” (Clotet/Torra 2010, 66).

By the end of the 1950s, the censors, more active than ever because there were more texts to be revised than ever, would see fit to allow the publication of a few periodicals, albeit only from ecclesiastical circles, which came under the protection of the Concordat. Thus, in 1958, the Benedictine monastery of Montserrat released issues of *Qüestions de Vida Cristiana*. The following year, also from Montserrat, the internal gazette *Germinabit* became *Serra d'Or*, a magazine which had an initial print run of 8,000 copies – a huge number for a magazine of high culture anywhere in Europe. Also, in 1959, a philosophy magazine called *Criterion* was published by monks of the Capuchin order. The blanket protection of the Church led to the publication, in 1961, of a children's magazine, *Cavall Fort* (with a circulation of 40,000 copies and 6,000 subscribers), and then, for adolescents, *Oriflama*. With these works in their hands, a minority of young people could see, for the first time in their lives, the language they spoke at home written with dignity; it was made visible as a distinct and respectable language. The Catalan language, moreover, could also be *spoken* with pride. In 1959, the seed of what would become the foremost phenomenon of the masses in Catalan would be planted: “La Nova Cançó”, which would resonate for the best part of a decade throughout the Catalan-speaking territories.

3 Anomalous normality (1960–1975)

At the beginning of the 1960s, Catalan under the Franco regime entered a third phase. In debates held in 1961, the highest consultative body of the dictatorship recognized the futility of the policies employed against the use of Catalan since the end of the Civil War. Professor Martí de Riquer declared in his *ad hoc* report, that Catalan could not be eradicated, nor was it possible to conceal its existence. The only solution to the problem was to treat the Catalan language as “una realidad idiomática Española”. To this end, Riquer (an adherent of the regime from the very start) proposed that the state consider the language of the Catalans as something of their own; that it be introduced at universities as a separate course, that optional classes be offered in secondary schools, that it be allowed in the cultural pages in newspapers and literary

magazines and that the *Jocs Florals* and the Institut d'Estudis Catalans be legalized. Riquer also suggested that the regime try to attract Catalan intellectuals to their cause (Santacana 2000, 43–47).

Riquer had in mind the recent centenary celebrations of the poet Maragall (with the participation of the head of state himself) and, especially, the award of the state's highest cultural distinction to the Catalan journalist, dramatist, novelist and poet Josep M. de Sagarra. Sagarra was presented with the award despite the fact that, in 1960, he had signed the manifesto, *Per la llengua catalana*, along with 100 other luminaries, in which they demanded, in accordance with UNESCO doctrine, 'regular school lessons' ('classes regulars') in the historical language of the country, and that 'the normal and recommended facilities be provided for the official functioning of the scientific and literary bodies that serve to cultivate and disseminate Catalan culture in its highest expressions' ('donades les normals i recomanables facilitats al funcionament reglamentari de les entitats científiques i literàries que tenen per objecte el conreu i la difusió de la cultura catalana en les seves expressions més elevades', Rafanell 2011, 466–467). In 1963, around 10,000 signatures were sent to the head of the government to demand the normalized use of Catalan in schools, the press and public administration. All of these initiatives were met with silence from the authorities.

However, in 1964, the regime proposed a series of gestures recognizing Catalan. A very small part of the official literature celebrating the 25 years since the founding of the Franco government was disseminated in Catalan. A minister even recited items of Catalan verse on the radio. This was the very first time since the Second Republic that the Catalan language had been used for official government purposes, albeit only briefly (Rafanell 2018). Following the creation of a new pre-university course, Catalan novels were included in the literature subject lists. At the end of the year in which Antoni M. Badia i Margarit published his well-known work, *Llengua i cultura als Països Catalans* (1964), a Chair in Catalan Philology was created at the University of Barcelona (which was about to be the embryo of a future Department of Catalan Philology). Spanish television, controlled by the state monopoly, broadcast a work by Josep M. de Sagarra in Catalan, and opened up a tiny doorway to the theatre (once a month). All this gave the impression of a change of attitude; nevertheless, in reality, the fundamental premise of imposed monolingualism remained rock solid.

The most decisive impact on the visibility of Catalan as an unregulated, but admissible language came from the environs of the Church. In this mostly church-going country, there were still reverberations from the papal encyclical, *Pacem in Terris*, issued by Pope John XXIII (1963) in which the rights of "minor nations" were defended; 200,000 copies were distributed in Catalan. Shortly afterwards, the reforms of the Second Vatican Council led to the introduction of the use of "vernacular languages" in the liturgy and the sacraments. When the church's new rules were implemented, Catalan – banned in official schools and in the mass media – regained a hugely significant distinction: co-officiality. The socialization of the corpus of language drawn up before the Civil War was given a huge boost when the Latin mass

was abandoned and the sermons and the liturgy were given in Catalan, regulated by the II Congrés Litúrgic de Montserrat (1965). The missals in standardized Catalan that had first been issued in 1964 (the classic *Foment de Pietat* and another by the monks of Montserrat, published in Andorra) channeled a whole series of normative standards (written and oral) which, until then, had only been expressed in literature. In the Balearic Islands, the Catalan liturgy was immediately adopted, with slight modifications that reflected the local dialect. In Valencia, however, a Vatican council missal in the local language (preceded symbolically by the 1951 *Eucologi valencià*) would not appear until 1975.

The Franco regime, which was highly sensitive to developments in the Catholic Church, continued to open up to the world and, in 1966, this led to the end of the previous linguistic censorship. Even before this, in 1962, as a new generation of Franco's politicians arrived on the scene, several publishers had already been allowed not only to use a now well-established standard Catalan, but also to include modern themes in their output. In 1964, Edicions 62 published *Els altres catalans*, by Francesc Candel, which tackled the problem of integration – including linguistic integration – among the huge numbers of immigrants flowing into Catalonia from southern Spain. The population of Catalonia in 1960 was just under 4 million. Between then and 1975 this figure grew by more than 1,770,000 of which about 950,000 were Spanish-speaking immigrants. At the end of 1964, Antoni M. Badia i Margarit carried out a survey to determine the extent of Catalan language usage among families in the city of Barcelona. The sample was small but representative, and showed – even then – that 62% of the population were Catalan-speaking.

Less than ten years had passed since translations of contemporary authors into Catalan had been prohibited. Now permission to do so was granted to a group of Barcelona publishers who until then had worked only in Spanish. Juventud brought out the adventures of *Tintin* in Catalan; Bruguera set out to publish 200 translations of juvenile literature. The prestigious Alfaguara publishing house issued a collection of dime novels in Catalan. Edicions 62 launched a collection of detective novels *La Cua de Palla* where police and criminals expressed themselves in Fabra's Catalan. Aymà published Fleming's *James Bond* in Catalan. The Proa publishing house, lost since the times of the Republic, reappeared from exile. In 1966, with the new law on press freedom in effect, the weekly magazine *Tele/Estel* began selling in Barcelona; it was the first periodical printed entirely in Catalan since the end of the Civil War. 1968 saw the return of the most successful of the pre-war children's magazines, *Patufet*, with a print run of 100,000. And while newspapers in Catalan remained unauthorized, some of the Spanish-language papers provided supplements in Catalan.

The growth in Catalan publications brought increasing respectability to the corpus throughout the Catalan-speaking territories. In Majorca, the linguist Francesc de Borja Moll continued his work of editing and, in 1962, he published the final volume of the *Diccionari català-valencià-balear*. In 1968, the magazine *Lluc*, a revival of the formerly bilingual *Lluch*, was published entirely in standard Catalan. In Valencia, where

linguistic particularism remained strong, all the post-war cultural products used a language that converged with general Catalan, except for some exotic cases or sub-literature published in the fiestas of *Les Falles*. In 1949, Editorial Torre had once again adopted the 1932 *Normes de Castelló*, taking these rules further, in order to capture an audience that went beyond Valencia. The publications of the long-established cultural society, Lo Rat Penat, did the same until the end of the dictatorship. However, it was not until the 1960s that publishers such as L'Estel (1962), Garbí (1967), *Gorg* (1969–1972) and, above all, *Tres i Quatre* (1968) connected the parallel-standard Valencian to modern content, as well as to translations into modern Valencian.

Meanwhile, official recognition of the language remained remote from the growing social demands of the day – but not entirely. In 1967, the Barcelona Provincial Council held its first examinations in Catalan to fill teaching posts in municipal schools, which were to be allowed one hour of Catalan classes a week. In August 1970, the Educational Reform Law was passed, promoted by the Valencian Minister José Luis Villar Palasí. It was the first big overhaul of education in Spain since the *Ley Moyano* of 1857. This time, ‘native languages’ were introduced in the educational programs, albeit through the back door. More than 1,500 Catalan corporations had demanded this change from the Cortes. According to the final disposition of the law, Catalan classes could only be optional; they had to be taught outside school hours, and would depend on demand. Five years after the new law had been passed, the proportion of children in Barcelona who had chosen the one hour of Catalan had not reached 10 %. Perhaps given the ‘uselessness’ of the subject, this was in fact quite a large number.

On May 30, 1974, General Franco, ill and decrepit, signed a decree from the Ministry of Education and Science governing the optional education of “las lenguas nativas españolas”. In 1975, the dictatorship ended, but it took three years before the mandatory introduction of the Catalan language into study programs at all levels was established by law. All the democratic political parties demanded that the official status of Catalan – abolished during the dictatorship – be restored. Safeguarded by the Spanish Constitution (1978), Catalonia’s 1979 Statute of Autonomy established Catalan as its ‘own language’, giving it co-official status with Spanish, which remained the only official language in the Spanish state as a whole. There soon followed similar legislation in Valencia (1982) and the Balearic Islands (1983).

4 The adoption of standards

The abnormal situation imposed on Catalan by the Franco dictatorship was not conducive to any revision of the corpus of the language. One of the first books authorized after the post-war persecution was the re-issue, in a single volume, of the encyclopedic dictionary *Pal·las* (1947), by Emili Vallès. Between 1951 and 1952, both Antoni M. Badia i Margarit and Francesc de Borja Moll published (in Spanish, of

course) historical grammars of the Catalan language. In 1954, the second edition of Fabra's *Diccionari general* was published. In the prologue, the poet Carles Riba, acting president of the Philological Section of the IEC, declared that 'the Fabra' was not a simple sketch of common Catalan, but a fundamental basis of future growth motivated by the work of philologists and literate men. Fabra's posthumous grammar book, edited by Joan Coromines in 1956, marked another milestone in the process of re-affirming the work done before the war, in this case open to the plurality of regulatory solutions, corroborated by Coromines himself.

In parallel with the reappearance of Fabra's major works, in 1950, the third volume of the *DCVB* had appeared. Francesc de B. Moll, who on the death of Antoni M. Alcover had taken on the task of editing the dictionary (assisted by the Valencian Manuel Sanchis Guarner), sought validation of the work, arguing that his dictionary aimed also to become normative, and compared favorably to the necessarily limited criteria of the *Diccionari Fabra*. However, the problem for writers as to which standards to adopt emerged above all with the visible increase in the number of texts allowed in Catalan between the mid-1950s and the beginning of the 1960s. At that time, with more or less stable periodicals and high-quality books on the market, the logical lexical and grammatical restrictions to which all writing was subjected by Fabra's norms generated all kinds of debate which, though rarely aired publicly, revealed a latent mood of inconsistency in the world of Catalan letters.

In 1955, Joan Sales, a former Catalan teacher returned from exile, founded the Club dels Novel·listes. The works published by the Club during the years of the publishing revival reflect a preference for the living language over the rigidity of the standardized language. In the foreword to *Bearn*, the Majorcan novel by Llorenç Villalonga, Sales justified the presence of non-standard language: 'a crime that is not forgiven by those who are self-righteously scandalized by the very expression 'living language' – perhaps because they would prefer it dead or mummified' ("crim que no perdonen aquells a qui la sola expressió 'llengua viva' escandalitza farisaicament, qui sap si perquè la voldrien morta o momificada", Sales 1961, 12). Sales used identical criteria in publishing *La plaça del Diamant* (1962), by Mercè Rodoreda, one of the great Catalan novels of all time, awash with words that did not appear in the official dictionary (Rafanell 1993). On the cover of *L'espera* by Sebastià Juan Arbó, the first novel to include the dialect of Tortosa in the dialogue, the publisher warned that 'we will perhaps have to endure the inconsolable tearing of tunics as so many ordinary people [...] do not swear according to the standard grammar and dictionary' ("potser assistirem a l'esquinçada de la manta túnica inconsútil, inconsolable perquè tanta gent del món [...] no reneguen conforme a la gramàtica i el diccionari normatiu", Sales 1967, 15–16).

Ramon Aramon i Serra, secretary of the IEC and the driving force behind its survival, was in favor of maintaining the Fabrian order against those who challenged it, to whatever extent; even those who simply allowed themselves a certain degree of flexibility. In a 1957 paper, he decried the syntactic deviations of contemporary story-

tellers (Aramon i Serra 1997, 721–755). In 1960, Aina Moll, who worked with her father on the *DCVB*, spoke of two opposing positions: those who believed that ‘they themselves are responsible for their work and therefore their form of expression must be respected’ (“que els responsables de la seva obra són ells, i per tant la seva forma d’expressió ha de ser respectada”) and those who believed that ‘in the current situation of our language, rigorous discipline needs to be applied in order to avoid degeneration’ (“en la situació actual de la nostra llengua cal sotmetre’s a una fèrria disciplina per evitar la degeneració”, Pericay 1987, 108). The second group could rely on the often unconditional adherence to the norms of publishing editors, trained in the old school. One of the most influential of these was Eduard Artells, former collaborator of Fabra and author of a *Vocabulari català-castellà abreujat* (1958), and vice versa (1961), and a language editor at *Serra d’Or* from its beginnings.

For Artells, the correct attitude of the Catalan proof reader consisted of ‘applying the grammar and safeguarding the purity of the language while looking forward to a brighter future in which his work could be reduced to that of correcting typing errors’ (“aplicar la gramàtica i vetllar per la puresa del llenguatge tot esperant temps millors en què la seva labor pugui restar reduïda a la d’estricta corrector tipogràfic”, Artells 1963, 123). This criterion, however, was often detrimental to much original work. For example, 1965 saw the publication of the second edition of *Vida privada* (1932), the great Barcelona novel by Josep M. de Sagarra, with an array of changes justifiable only by the scruples of a purist. In 1968, the Fundació Bíblica Catalana updated in a single volume the biblical texts that it had previously published between 1928 and 1936. The ‘correct Catalan’ which, from 1961 onwards, was learned by candidates for teaching posts in the private institution Òmnium Cultural, was basically the Catalan derived from pre-war standardized Catalan. It was reinforced in the early 1960s by two prestigious publishing houses. First, there were the fascicles that would eventually comprise the four volumes of the *Diccionari enciclopèdic Salvat* (1968) – a symbolic reprisal of the series published between 1930 and 1935 by the same company – under the direction of Albert Jané. Second, there were the 15 volumes of the *Gran enciclopèdia catalana* (1969–1980), under the linguistic direction of Ramon Aramon i Serra. These two works were fundamental in the socialization of standard Catalan.

In 1968, the centenary of the birth of Pompeu Fabra was commemorated. This was an important year in the restoration of the thesis of the grammarian. Ramon Aramon re-edited Fabra’s school grammar of 1918. Eduard Artells revised and extended the *Curs pràctic de gramàtica catalana*, by Jeroni Marvà (pseudonym of Artur Martorell and Emili Vallès) which had appeared between 1934 and 1937. Re-titled *Curs superior de gramàtica catalana*, it became a reference text for learning Catalan in the final years of the Franco period. Even so, there were voices at that time that began to question the old norms, and their pedagogical suitability. Joan Solà (1973), in the first international congress of Catalan Studies, celebrated in Strasbourg in 1968, openly defended reforms which he would later develop in a considerable body of work

(*Estudis de sintaxi*, 1970; *Del català incorrecte al català correcte*, 1977; *Qüestions controvertides de sintaxi catalana*, 1987; and others). Gabriel Ferrater was critical of the standards in his column “De causis linguae” in *Serra d’Or* (1968–1972). The teacher Lluís López del Castillo, for his part, shed light on the utilitarian polymorphism of the language in *Llengua standard i nivells de llenguatge* (1976). The essayist Joan Fuster, who always welcomed the usefulness of a “national” model for those of his texts not specifically aimed at Valencians, reproached the old correctors for their ‘language excesses and grammatical excesses’ (“excessos de llenguatge, excessos de gramàtica”, Pericay 1987, 154–158).

More than anything, the shakeup of many of the hitherto undisputed normative certainties was mainly the result of the penetration of Catalan in the mass media, the press, on the radio and on TV. With Franco dead, a Catalan-language newspaper, *Avui*, began publication in 1976, eventually followed by others (*Punt Diari*, *Diari de Barcelona* as well as Catalan versions of *El Periódico* and *La Vanguardia*). In 1983, the television channel TV3 was launched, along with the radio station Catalunya Ràdio. These new entities brought with them no end of ‘style guides’ (“llibres d’estil”), each aiming to harmonize the language of the reference books with the various language registers in existence, and to combine the genuineness of the corpus with the more true-to-life and intelligible reality of the language (↗21 Catalan in the Mass Media: The Rise of Stylebooks).

The process was not free of controversy. The theory behind the proposed changes was set out in the essay *Verinosa lengua* by Xavier Pericay and Ferran Toutain (1986), who followed this up with *El malentès del Noucentisme* (1996), a critique of the traditional expressive options. Many users of the formal language saw ‘no need to forgo the option of saying or writing, in more or less formal language, certain manifestations of what is understood as real, genuine and generally-used language’ (“necessitat de no poder renunciar a dir o escriure en llenguatge més o menys formal certes manifestacions del que s’entén per llengua real genuïna i d’àmbit general”) and, in this way, they validated many consolidated living forms which had, thus far, been excluded from the norms (Fité Labaila 2015, 49). The case of Quim Monzó, who in 1999 rewrote (and revised linguistically) all of his previous narratives, exemplifies the change of perspectives followed by many writers, journalists, translators and editors.

The decisive influence of the *Diccionari etimològic i complementari de la llengua catalana*, by Joan Coromines (1980–1995), led to the appearance of the IEC’s official *Diccionari de la llengua catalana* (DIEC) in 1995, with almost 16,000 more articles than the 1932 normative edition. 2,422 more entries were added to the second edition of the DIEC in 2007. As for the grammar, in 1994, Antoni M. Badia i Margarit published his *Gramàtica de la llengua catalana* with the aim of integrating the descriptive with the standard, as well as the diatopic and diastratic variants. The three volumes of the *Gramàtica del català contemporani* (2002), edited by Joan Solà, constituted the basis of what finally culminated in the 2016 version of *Gramàtica de la llengua catalana*, a collaborative work published by the IEC, which aimed to integrate the solutions put

into practice by the media but generally neglected by the official grammar, which was, until then, still Fabra's from 1918.

5 Between rupture and unity

The Franco regime never really challenged the idea of Catalan language unity because it saw no need to do so. The anti-Catalan attacks in the Majorcan press in the final years of the Franco regime were fairly insignificant and lacking in tradition (Moll 1972). The inveterate local sensibilities of the people of Valencia would have been more conducive to such arguments, but were a long time coming. The language courses promoted by Lo Rat Penat since 1949, tolerated and endorsed by the Regime, had always been inclusive of the entire Catalan-speaking area in terms of the norms and geographical aspects being taught. The text book used, *Gramàtica* by Carles Salvador (1951), was designed to elucidate a grammar by Sanchis Guarner (1950) which openly converged with general Catalan. In 1951, work had resumed on the *DCVB*, new instalments of which were presented to the Valencia City Council. With the backing of the Franco authorities and announcements on official radio, the event drew large crowds. It would appear that the idea of 'the Catalan danger' ("perill català") – the title of a provocative, pre-war essay by Josep M. Bayarri – had receded into the distance. There was no 'danger', because Catalan presented no problems in a Valencia where the historical language was functionally subordinate. Even so, in the late 1950s, writer and librarian Nicolau Primitiu Gómez Serrano attempted to revive a proposal he had made thirty years earlier: to use the acrostic *Bacavès* as a substitute for Balearic, Catalan and Valencian. On the other hand, in the magazine *Sicània*, he also continued to argue that Valencian had originated independently of Catalan, an idea linked to the old theories of Lluís Fullana. For those able to think this way, separate historical origins were no obstacle to accepting the undoubted present-day unity.

Valencian idiomatic secessionism did not take hold of society to any significant extent until the early 1960s. In 1960, Manuel Sanchis Guarner published a completely rewritten edition of his book for schoolchildren, *La llengua dels valencians*. However, the work that had the strongest impact on Valencian consciousness was the essay *Nosaltres, els valencians* (1962), written by an intellectual held in high regard throughout the Catalan-speaking territories, Joan Fuster. The assumption of a unitary language was confirmed in the dictionaries by Francesc Ferrer Pastor (1960, 1966) and in the *Curso de lengua valenciana* by Enric Valor (1966). However, towards the end of the Franco regime, while the conservative Valencianism of Lo Rat Penat continued to publish the unitarian grammar of Carles Salvador, the conflict over the identity of the Valencian language took a violent turn. In 1974 – with the decisive influence of the mass media – a real battle began in Valencia, the main objective of which was to stop the supposed "catalanization" of "Valencian" and, as a logical consequence, of Valencia itself. Apart from the name of the language itself, the conservative press also

revived the archaic interpretation, based on a presumed ‘Mozarabic origin’ (“mossarabisme”), which considered Valencian and Catalan to be of different origins. In 1977, Miquel Adlert and Xavier Casp, who had founded the publishing house Torre shortly after the Civil War, led intellectual attempts to refute any Catalan foundations of the Valencian language. The following year they spearheaded the creation of the Acadèmia de Cultura Valenciana, overseen by the Provincial Council of Valencia, which would create an alternative and openly separatist orthography.

In 1978, the Faculty of Philology at the University of Valencia approved a report entitled *Informe sobre la llengua del País Valencià*, which detailed the origin of the language, the solution to the problem of its name and the adoption of standards for Valencian, based on criteria of polycentrism and convergence with general Catalan. In 1982, the Statute of Autonomy of what was henceforth to be called the “Comunitat Valenciana”, decreed the co-official status of “Valencian” but left the question of identity unresolved. The political to-and-fro during the 1970s and 1980s kept the discord alive between a secessionist right wing and a Catalan-leaning left wing. It was not until 1998 that, with the creation of the Acadèmia Valenciana de la Llengua, the Valencian government established an organ competent in the regulation of the language. The *Diccionari ortogràfic i de pronunciació del valencià* (2006), the *Gramàtica normativa valenciana* (2006) and the *Diccionari normatiu valencià* (2014) brought the discord to an end, proposing standards for Valencian that, while recognizing the unity with the language elsewhere known as Catalan, conformed to autonomous rules and resisted submission to the dictates of any kind of outside regulation.

6 Bibliography

- Aramon i Serra, Ramon (1997), *Estudis de llengua i literatura*, Barcelona, Institut d'Estudis Catalans.
- Artells, Eduard (1963), *Actitud correcta del corrector*, Criterion 17, 117–123.
- Ballester, Josep (1992), *Temps de quarantena. Cultura i societat a la postguerra (1939–1959)*, València, Tres i Quatre.
- Barba, Bartolomé (1948), *Dos años al frente del Gobierno Civil*, Madrid, Javier Morata.
- Benet, Josep (²1978, ¹1973), *Catalunya sota el règim franquista. Informe sobre la persecució de la llengua i la cultura de Catalunya pel règim del general Franco (1a part)*, 1a reedició, Barcelona, Blume.
- Berjoan, Nicolas (2011), *L'identité du Roussillon: Penser un pays catalan à l'âge des nations (1780–2000)*, Perpignan, Trabucaire.
- Clotet, Jaume/Torra, Quim (edd.) (2010), *Les millors obres de la literatura catalana (comentades pel censor)*, Barcelona, Acontravent.
- Fité Labaila, Ricard (2015), *Quan l'ús és la norma*, in: Magí Camps et al. (edd.), *Canvi d'agulles*, Barcelona, RBA, 46–61.
- Gallofré i Virgili, Maria Josepa (1991), *L'edició catalana i la censura franquista (1939–1951)*, Barcelona, Publicacions de l'Abadia de Montserrat.
- LV = *La Vanguardia Española* (1881–).

- Manent, Albert/Crexell, Joan (1988), *Bibliografia catalana dels anys més difícils (1939–1943)*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Marín i Corbera, Martí (1996), *Existí un catalanisme franquista?*, in: Pere Anguera et al. (edd.), *El catalanisme conservador*, Girona, Cercle d'Estudis Històrics i Socials, 271–292.
- Menéndez-Reigada, Albino (2003), *Catecismo patriótico español*, ed. Hilari Raguer, Barcelona, Península.
- Moll, Francesc de B. (1972), *Polèmica d'en Pep Gonella*, Palma de Mallorca, Moll.
- Pérez i Vallverdú, Eulàlia (ed.) (2009), *Fantasmes rojos. La venjança falangista contra Catalunya (1939–1940)*, Barcelona, Acontravent.
- Pericay, Xavier (ed.) (1987), *L'altra cara de la llengua*, Barcelona, Empúries.
- Rafanell, August (1993), *La llengua inadvertida de Mercè Rodoreda*, Revista de Girona 157, 60–63.
- Rafanell, August (2011), *Notícies d'abans d'ahir. Llengua i cultura catalanes al segle XX*, Barcelona, Acontravent.
- Rafanell, August (2018), *Una rectificació del franquisme sobre el català*, in: Carles Santacana (coord.), *Quan tot semblava possible... Els fonaments del canvi cultural a Espanya (1960–1975)*, València, Universitat de València, 177–205.
- [Sales, Joan] (1961), *Nota dels Editors*, in: Llorenç Villalonga, *Bearn*, Barcelona, Club Editor, 9–12.
- Sales, Joan (1967), *De l'editor a l'autor*, in: Sebastià Juan Arbó, *L'espera*, Barcelona, Club Editor, 13–24.
- Santacana, Carles (2000), *El franquisme i els catalans. Els informes del Consejo Nacional del Movimiento (1962–1971)*, Catarroja/Barcelona/Palma, Afers.
- Solà, Joan (1973), *Ortographe et grammaire catalanes*, in: Antonio Badía Margarit/Georges Straka (edd.), *La linguistique catalane. Colloque international organisé par le Centre de Philologie et de Littératures Romanes de l'Université de Strasbourg du 23 au 27 avril 1968*, Paris, Klincksieck, 81–100.
- Solé i Sabaté, Josep M./Villarroya, Joan (1993), *Cronologia de la repressió de la llengua i la cultura catalanes (1936–1975)*, Barcelona, Curial.

Albert Turull

15 Onomastics: Personal Names and Place Names

Abstract: Proper names, the object of study of onomastics, provide unique insights into the history, territory and society in which they were generated and in which they are now used. It is this interdisciplinary nature of onomastics (including, in particular, toponymy and anthroponymy, which concern themselves with the proper names of places and persons, respectively) that makes it a complementary subject for many branches of study, as well as undoubtedly for linguistics. The different typologies of forenames and surnames in use within the territories of the Catalan language are a reflection of migrations, both ancient and modern; of religions and occupations, some still practised and others lost in time; and, of shifting fashions, and yet, at the same time, of a certain continuity over the centuries thanks to a process of intergenerational transmission. Similarly, the toponyms of Catalonia, while being a faithful reflection of its territory, constitute a living document of the various language strata that have been laid out across this area, and of many linguistic features that have fallen into disuse.

Keywords: proper names, place names, personal names, toponymy, anthroponymy

1 Catalan onomastics

Onomastics emerged as a discipline in its own right in the 19th century (coinciding in Catalonia with the Romantic revivalist movement known as the *Renaixença*: 1833 *Renaixença*), and over the course of the 20th century it enjoyed considerable development. Today, it is learning to adapt to the challenges posed by the information society and it enjoys a presence in the academic world as well as in the offices of regional planning and in the work of language consultancy services. From the outset, onomastics has been understood as a complementary, interdisciplinary subject, with geographers and historians making some of the earliest contributions. Significantly, the key work of the period, *Orígenes històrics de Catalunya* (Balari Jovany 1899), despite its obviously historicist approach, is the work of a philologist trained in the classics.

A qualitative leap was made during the 20th century, coinciding with the founding of the Institut d'Estudis Catalans (IEC). As part of its normative mission, the academy of the Catalan language showed an early interest in onomastics, creating a specific office in 1922 and publishing such works as the study conducted by the expert in Romance languages, Paul Aebischer, in 1926. Moreover, it was within the Institut that the philologist, Joan Coromines, who was to become the leading figure in Catalan

<https://doi.org/10.1515/9783110450408-023>

onomastics, began his career. His *opus magnum*, *Onomasticon Cataloniae* (*OnCat*), is an indispensable reference work for deciphering Catalan toponyms. Among his other publications – that is, onomastic data catalogued within an etymological dictionary – are the works collected under the general title of *Estudis de toponímia catalana*.

The work of Coromines was of enormous importance, but we should not forget the contributions of other scholars, above all those of Francesc de B. Moll and Enric Moreu-Rey. Moll was to play a key role in the continuation and completion of the *Diccionari català-valencià-balear* (*DCVB*), a monumental dictionary begun by Antoni M. Alcover which also incorporates a multitude of references to names both ancient and modern, while his work *Els llinatges catalans* (Moll 1982) is indispensable for studying surnames. As for Moreu-Rey, his work was to prove decisive in the dissemination of onomastics, thanks to the foundation of the *Societat d'Onomàstica* in 1980, which has brought together hundreds of scholars and lovers of the subject, and to the impact of his own publications, which cover all branches of onomastics, from toponymy (Moreu-Rey 1982) to anthroponymy (Moreu-Rey 1991), as well as nicknames and the names given to family homesteads (Moreu-Rey 1981) and even urban toponymy and miscellaneous onomastics, which he referred to as *onomàstica vària* (Moreu-Rey 1974).

2 Catalan toponymy

Catalan toponymy can be approached from three points of view, none of which are mutually exclusive; on the contrary, the perspectives overlap and bolster each other. These approaches can be described as looking at the following: the historical language stratum in which each toponym appeared; the toponym's semantic referent; and, finally, the morphological formation to which the toponym corresponds.

2.1 The historico-linguistic strata of Catalan toponymy

The Catalan language is the result not solely of the evolution of Vulgar Latin, but also of the contribution of various languages which, both before and after Latin, have been spoken in the Catalan territories. As such, Catalan toponymy has the value of a living document, while the study of the ancient toponyms of the Catalan domain can make notable contributions to our knowledge of these past languages.

2.1.1 Pre-Roman toponymy (Indo-European and non-Indo-European)

The most remote strata from which toponyms in use today have reached us are those that preceded Romanisation. Relatively little is known about this period (cf. Velaza et

al. 2002), but what seems certain is that there were at least two main substrates: the Indo-European and non-Indo-European.

The latter corresponds specifically to the Iberian culture and language, but any investigation today runs up against an insurmountable difficulty: the fact that the language cannot be deciphered. Moreover, the old theory of Basque Iberianism, according to which the Basque language or *euskera* was believed to be the sole remnant of an ancient Iberian that had spread throughout much of the Peninsula, has been completely discarded. Thus, we have to limit ourselves to identifying just a few names, which by way of deduction (and, on occasion, with the help of historical and even numismatic documentation) can be considered Iberian in origin. Among them, we find those of the capital cities of *Barcelona*, *Tarragona*, *Lleida* and *Tortosa*, whose present-day names derive from the base names of BARKENO, TARRAKO, ILTIRTA (or ILERSDA) and DERTOSA, respectively, via the Latin names of *Barcinona*, *Tarracona*, *Ilerda*, and the Arabized *Turtusa*. The phenomenon observed in the first two cases is systematic: the addition in Latin of the suffix -ONA, which affects a whole series of toponyms, including *Isona* < AESO, *Guissona* < IESSO, *Solsona* < CELSA, *Badalona* < BAETULO, and *Osona* < AUSA. A similar base is found without the suffix in the case of *Tàrraga* (< TARRAKA), and the origins of some major place names in the Valencian Community are also Iberian, and include *Alacant* < LEUKANTO, *Elx* < ILICI, and *Xàtiva* < SAETABIS. More recently the Iberian origins of *Segarra* (< SIKARRA) have been confirmed, and a concentration of such names has been observed in the north-east of Catalonia, including *Tossa* < TURISSA, *Guixols* < KUIKSALOS, *Aro* < AZARE, *Cotlliure* < CAUCO-ILIBERIS, and *Begur* < BEKURI.

The Basque language is not only pre-Roman but also pre-Indo-European, or at least non-Indo-European, and this – although we can rule out any connection with ancient Iberian – provides the key for deciphering toponyms, especially if we accept the theory forwarded by Coromines (1965, 93–152) that its historical reach was far more extensive, including the Catalan *Alt Pirineu* (or the High Pyrenees). Whatever the case, in order to distinguish these toponyms from what are properly speaking Basque place names, recourse is usually made to the term *Bascoide* (Coromines 1965, 153–217), revealing that a good number of names can be explained in relation to this stratum, above all those in the Pyrenean region: e.g. *Suert*, *Durro*, *Taiüll*, *Esterri*, *Unarre*, *Isil*, *Sorpe*, *Sort*, *Gerri*, *Ison*, *Boabi*, *Asnurri*, and *Estavar*, to mention just those from the central districts of the mountain range, although similar hypotheses can be made for toponyms from other regions (e.g. *Ivars*, *Gerb*, *Ivorra*, *Sanaüja*, and *Biosca*).

The Indo-European substrate only has the fact that it is pre-Latin (or pre-Roman), and a certain overlapping chronology, in common with the above. Yet, it is not easy to distinguish within this vast Indo-European substrate which cultures have left a trace in Catalan toponymy. Even Coromines, despite his well-known (but poorly accepted) theory of *Urnfelder*, or Urnfield, is unable to confirm hypotheses for pre-Roman names such as *Reus*, *Tivenys*, *Blanes*, those of the rivers *Isàvena*, *Cinca*, *Segre*, *Xúquer*,

Ter and *Tet*, and, in particular, the set of place names bearing the suffix *-esa*, which includes such relevant toponyms as *Manresa*, *Olesa*, *Artesa*, *Orpesa*, and *Montesa*.

The pre-Roman presence of the most clearly recognisable Indo-European type is that of the Celtic cultures, given the relative proximity of important settlements of this ethnic group. The origins of the following names have all been proposed as being Celtic: *Berga*, *Besalú*, *Onyar* and *Talavera*, with a marked concentration in given areas, especially the Conca de Tremp, where we find *Talam*, *Segú*, *Salàs*, and *Tremp* itself.

To finish, there are toponyms related to cultures that, unlike the above groups, were never widely established in the territory. These include a few names of Punic origin in the Balearic Islands (*Eivissa*, *Tagomago*, *Maó*), and some of Greek origin, a testimony to Hellenic colonies that grew up along the Mediterranean coast. The following toponyms are Greek in origin: *Gandia* < KANDIA ‘isle of Crete’, *Elna* < HELENA (anthroponym), *Empúries* < EMPORION ‘market’, and *Roses* < RHODAS ‘isle of Rhodes’.

2.1.2 Latin toponymy

Given the difficulty in distinguishing Latin toponyms from those that were formed directly in Catalan but based on common words of Latin etymology, this section only includes toponyms actually documented in Roman times (e.g. the capital cities of *València* < VALENTIA and *Girona* < GERUNDA) and those that present characteristics of the Latin language that are absent or unusual in the Romance languages and in Catalan in particular: e.g. *Sils* < SILVIS ‘in the woods’, *Traiguera* < TRITICARIA ‘fields of wheat’, and the river *Flamisel·l* < FLUMENCELLUM (diminutive of FLUMEN ‘river’).

Many names originate from either simple or derivative nouns: *Fraga* < FRAGA ‘gorge’, *Prada* < PRATA ‘meadows’, *Flix* < FLEXUM ‘meander’, *Vic* < VICUM ‘village’, *Òpol* < OPPIDUM ‘fortification’, *Sallagosa* < SALICOSAM ‘willow grove’, *Ceret* < CERA-SETUM ‘cherry trees’, etc. Similarly, a fair few are derived from adjectives (including participles): *Conesa* < [SILVA] CONDENSEA ‘dense [wood]’, *Amposta* < IMPOSITAM ‘sited [on the river bank]’, *Llobregat* < RUBRICATUM ‘ruddy [the colour of the river]’, *Guils* < EQUILES ‘meadows or stables for horses’, etc. Compound nouns are also common; e.g. *Morvedre* < MURUM VETEREM ‘ancient wall’, *Tírvia* < TRIVIA ‘three paths’, *Ribagorça* < RIPA CURTIA ‘cut riverbanks’, or *Igualada* < ACQUAM LATAM ‘broad water’ – indicating the point where the river *Anoia* (< AMNUCULAM, diminutive derived from AMNIS ‘water, river’) broadens out on its course down from the mountains.

But the most typical Latin toponyms are those derived from a personal name, either by reproducing it directly (e.g. *Constantí* < CONSTANTINUM) or, more frequently, by adding a diminutive suffix (e.g. *Argençola* < ARGENTEA) or a derivative form, indicating the founder (or owner) of the place: a villa belonging to Tiberius is TIBERIANAM > *Tiurana*, or a tower built by Granius, GRANIANAM > *Granyana*, *Granye-*

na. The most frequent device is the suffix -ANUM, and there are many examples, most of which were studied by Aebischer (1926) and confirmed, refined or added to by Coromines: some quite obvious, such as *Cornellà* < CORNELIUS, *Lluçà* < LUCIUS, and *Pallejà* < PALLADIUS, and others that have been made less evident owing to their subsequent phonetic evolution, such as *Juià* < JULIUS, *Teià* < TALIUS, *Avià* < AVITIUS, *Polinyà* < PAULINUS, *Premià* < PRIMILLUS, etc. Finally, the name of the capital of Northern Catalonia (today located in France and known as Perpignan), *Perpinyà*, shares this same origin, specifically from the Latin anthroponym PERPENNIUS.

2.1.3 Germanic toponymy

If, among the names of Latin origin, toponyms derived from personal names constitute an important group, this trait is even more important among the place names of Germanic origin, to the point that all the Catalan toponyms corresponding to this stratum in fact have an anthroponymic root. In Catalan toponymy of Germanic origin (the focus of particular study by Coromines 1965, 31–65), there are names that originate from both the Visigothic (5th–8th centuries) and the Carolingian (8th–10th centuries) periods – that is, before and after the beginning of Arab domination, respectively – meaning that they do not correspond to just one sole language, but rather to (a combination) of both Gothic and Frankish. This fact, coupled with the continuing predominance of the anthroponymy of Germanic origin from the 10th through to the 12th century (cf. Aebischer 1928), generated a large number of place names derived from these personal names.

In some cases, these are a direct reflection of a Germanic anthroponym (e.g. *Fals* < FALKS), while in others an inflective variant -AN (e.g. *Berà* < BERAN, *Guimerà* < WIGMERAN) is used. There are examples of other inflections and affixes (e.g. LOTFRI-DI > *Llofriú*), as well as toponyms that incorporate an anthroponymic element as a complement combined with a generic Romance element (e.g. VILLA ALIHARI > *Vilaller*, VILLA EDRAD > *Viladrau*, PALATIUS HRABAN > *Parlavà*). But the Germanic anthroponymic system typically resorted to compositions of two significant elements. Especially productive in this regard were, among many others (see the vast compendium compiled by Förstemann 1900), base forms such as HARI and NANTH. These gave rise to such Catalan toponyms as *Folquer* < FULKHARI, *Gualter* < WALTHARI, *Renan* < RAGINANTH, and *Senan* < SISENANTH, the latter sharing the same root as *Sineu*, although it subsequently underwent a different phonetic evolution. Better known is the transformation undergone by toponyms that contain the anthroponymic element SIND, which was transformed into -rèn or -reny by rhotacism: e.g. *Gombrèn* < GUMESIND, *Gisclareny* < GISCLASINN (among others that were studied by Coromines 1965, 32–39).

2.1.4 Arabic toponymy

The notable presence of an Arabic stratum in Catalan toponymy is the direct consequence of Arab dominance over a large part of the territory between the 8th century and, depending on the region, the 9th and 13th centuries. This implementation of such differing chronology is reflected in the degree of toponymic intensity that varies depending on the area in question (see Coromines 1965, 265–279). Thus, while in the *Alt Pirineu* Arabic toponymy is non-existent and scarce in the rest of *Catalunya Vella* (Old Catalonia), in *Catalunya Nova* (New Catalonia) – where the period of Muslim rule lasted until the middle of the 12th century – it is quite abundant, and even more so in the Balearic Islands and the Valencian Community, where the Arabs remained until the 13th century.

Unlike Germanic toponymy, that of Arab origin presents a wide variety of semantic referents and also of formal possibilities (cf. Barceló 2010). Thus, alongside simple nouns, often in conjunction with the fossilised article *al* (examples of which are given below), we also find compound nouns (e.g. *Burjassot* ‘black tower’, *Massalió* ‘hostel of the springs’), and compounds of a common noun and a personal name (e.g. *Margalef* ‘meadow of Halef’, *Sedaví* ‘lord [sidi] Dabbi’, *Sidamon* ‘lord Hammun’) and even references to ethnic or tribal origins (*Massalcoreig* ‘hostel of Kuraixi [Arabia]’).

Among the simple nouns (together with various adjectives: *Alcoi* ‘rich’, *Albaida* ‘white’) some refer to the natural environment (*Alzira* ‘island’, *Alcúdia* ‘hill’), others to agricultural crops (*Aitona* ‘olive trees’, *Alcarràs* ‘cherry tree’), although there are also references to military constructions (*Alcàsser* ‘fortress’, *Almenara* ‘tower’) and to religion (*Almacelles* ‘prayers’), without forgetting more specific architectonic elements (*Alaquàs* ‘arches’). This same multiplicity of referents – characteristic, in short, of a complete, well-established toponymic system – also occurs in the names that have not retained the Arabic article (or which would later adopt the corresponding Romance article: e.g. *la Nucia* ‘walk’, *la Ràpita* ‘monastery of warriors’). Again, we return to elements of geography with *Xeresa* ‘hard land’ and *Saidí* ‘arm of the river’; military architecture in *Miravet* ‘advanced fortification’, civilian architecture in *Simat* ‘quarter’ and *Masquefa* ‘house of play’; and a reference to a certain lordly domain in the name *Gallifa* ‘caliph’.

The last example brings us to names derived from anthroponyms, which are also frequent in the toponymy of Arabic origin. Some allude directly to a personal name (*Bràfim*, *Calaf*, *Ifac*, *Jafre*, *Mafumet*), but more common is the patronymic formation in which the *ibn* (‘son of’) is found fossilised in the form *Vin-* (e.g. *Vimbodí*, *Vinaixa*, *Vinalesa*, *Vinaròs*) or, more frequently – especially in the Valencia area – *Ben-* or *Beni-*: *Benissa*, *Benialí*, *Benicàssim*, *Benifaió*, etc.

2.1.5 Pre-Catalan Romance toponymy

The so-called *Mozarabic* (also referred to as *Andalusi Romance*, or, even, the pre-Catalan Romance languages) has also left its mark on the toponymy. Yet, it is difficult to distinguish the names attributable to this stratum from those of a strictly Latin origin, but there are some characteristic traits – generally of an archaic character – that allow us to identify these place names with some degree of certainty. Thus, among the sample of toponyms we find those that have typically conserved a final vowel (e.g. *Muro*, *Campello*, *Campos*), a feature that can also occur in the first element of a compound noun (*Portopi*) or, indeed, in both (*Portocristo*). Other distinguishing features include the non-sonorisation of occlusives (*Catí* < CATINUS ‘hollow’, *Petrer* < PETRARIUM ‘stony’) and unique phonetic evolutions (*Peníscola* < PAENINSULA ‘peninsula’, *Polop* < PLOPPUS < POPULUS ‘poplar’).

Significantly, a similar phenomenon – but one that cannot be attributed to an Arabic influence, but which is a pure archaism related also to a late process of Romanisation or Catalanisation – occurs in the High Pyrenees of Catalonia, where some of those linguistic features are found, including the very clear conservation of final vowels (*Monestero*, *Molinos*, *Avellanos*, *Sallente*) and the non-sonorisation of occlusives (*Llevata* < GLEBATA), among other features that are even more specific or subtle.

2.1.6 Toponymy of adstrate languages

We consider toponyms from the adstrate as those that were not generated directly in a territory, but which have been copied or transplanted from another country (usually neighbouring, in the case that concerns us here, although in the case of the Americas there are transoceanic examples resulting from the processes of conquest and colonisation).

In the case of Catalan, toponyms of foreign origin (with some notable exceptions, such as the Italian *Benevento* > *Benavent* and the Franco-Belgian *Aigremont* > *Agramunt*) are almost always of Occitan origin. This can be attributed to geographical contiguity and a certain cultural and linguistic proximity, but also to a historical period (10th–12th centuries) of territorial expansion. For this reason, in *Catalunya Vella* we find such toponyms as *Avinyó* and *Rodés*, and, in particular, a concentration of such place names in the area that stood at the frontier between the 10th and 12th centuries, with cases like *Verdú* and *Montbrío*.

2.1.7 Authentic Catalan toponymy

In spite of the remarkable diversity of language strata present in this territory, most of the toponyms in the Catalan domain derive, in terms of their linguistic origin, directly

from Catalan. This can be attributed to the fact that the consolidation of the territory's social structures, settlement patterns and farming occurred in parallel with the expansion, in the first instance, of the Catalan counties and, later, of the former kingdoms of Majorca and Valencia. In other words, the majority of place names were formed in a period in which the language spoken was without doubt Catalan, despite its coexistence with other languages.

Clearly, from a language perspective, certain toponyms such as *Mont-roig* and *Montblanc* must be considered as Catalan despite the Latin and Germanic etymology, respectively, of the adjectives *roig* 'red' and *blanc* 'white'. The same can be said for many other place names, a list that runs into the thousands if we include the minor toponymy. Thus, we will not extend this section much further with the reasoning that, in fact, will be developed in the following sections, where we break the Catalan toponymy down into its distinct semantic fields. Here, without straying outside the framework of the toponyms of Catalan etymology, we shall limit ourselves to pointing out that a series of highly interesting phenomena in terms of historical linguistics emerge from their study, emphasising once more the documentary value of onomastics: for example, the presence of lexical elements from mediaeval Catalan that have disappeared from the modern language, such as the variant *pena*, with /n/, from the common *penya* < lat. PINNAM 'rock' (present in the simple noun *la Pena*, the derived noun *Penelles*, and in the compound noun *Pena-roja* 'red rock'); and the adjective *fret*, derived from the Latin FRACTUM 'broken', easily confused with *fred* < FRIGIDUS 'cold' in a toponym such as *Montfred* and even with *fet* < FACTUM 'done' in *Torrefeta* (formerly *Torrefreta*). In a toponym like *Sapeira* < IPSA PETRA 'the rock', two phenomena stand out: on the one hand, a variant of the diphthong *peira*, which is found in other Catalan toponyms in the form *pera* (from the simple noun form *la Pera* to compound nouns such as *Peralta* 'high rock' and *Peratallada* 'cut rock'); and on the other, the archaic article (today also dialectal, primarily Balearic) *sa* < IPSAM, which survived for centuries alongside the more widespread *la* < ILLAM, but which, before disappearing from most of the Catalan territory, was fossilised in numerous place names throughout practically all the Catalan-speaking regions. Alongside *Sapeira* we find the more obvious *Sarroca* 'the rock', and examples in the plural also exist (with the feminine article *ses*, often accompanying a saint's name: *Sesoliveres*, *Sesrovires*, with the lexemes *olivera* 'olive tree' and *rovira* 'oak grove') as they naturally do in the masculine (lat. IPSUM), as in *Sant Joan Despí*, agglutination of *d'es pi* ('of the pine'), and *Espui* (*el puig* 'the hill'). Less obvious are the cases in which the article, preceding a vowel, takes the reduced *s*' form, so that it is a single initial /s/ in names that, on occasions, have been given other interpretations (e.g. *Solivella*, where *olivella* is a suffixation of *oliva* 'olive', that is to say, 'olive tree'; or *Suró*, where *uró* is actually a dialectal spelling of *oró*, and this a monophthongisation of *auró* < lat. ACERONEM < ACER 'maple').

2.2 The semantic fields of Catalan toponymy

In Catalan toponymy – in common with all toponymies – reference is made to a wide variety of motifs. Whether it be in Catalan itself, Latin or Arabic, or any other of the language substrates that coincide within the Catalan territories, place names allude (or in their origins alluded) to a series of referents, most of which can be described as being *objects*, insofar as they are readily identifiable elements of the surrounding environment. Geographical features, vegetation, fauna and buildings make up the main semantic fields that the toponymy draws on. Alongside these, however, we find other classes of names in which the referent responds to a more intangible cultural or historical circumstance (the ancient name of a person, a religious dedication, etc.). For this reason, we can speak of certain degrees of arbitrariness in their denomination, inversely proportional to their tangibility. In the following sections, we review these different types of place names, associated with varying degrees of rootedness in the environment.

2.2.1 Names referring to the physical environment

First, we find names associated with the natural environment, that is, to elements of the territory unrelated to its human occupation. This group is headed by toponyms referring to the local landscape and the shapes of the terrain, be it in a simple, straightforward manner (*el Puig* ‘hill’, *el Tossal* ‘hill with high plateau’), employing a diminutive suffix (*Pujol* ‘small hill’, *els Muntells* ‘small rises’), or by providing a descriptive composition (*Montserrat* ‘mountain shaped like a serrated saw’). In this group, we also find toponyms that employ a noun with a metaphorical function in reference to a specific orographic formation; e.g. *Jou* ‘yoke’, or *Escales*, *Escaló* and the like, which contain the word *escala* ‘stair’. Other names identify some characteristic of the quality of the terrain, such as its colour; *Claret*, like *Montclar* and, perhaps, *Vallclara*, are centred on the adjective *clar* ‘bright’. Equally well known are names linked with the Latin word *RUBEUS* ‘red’ – and its Romance derivatives – such as *Rubí*, *Rúbies*, *Rubió*, *Rojals*, and the compound names of *Font-rubí*, *Mont-roig*, etc. Similarly, the texture of the terrain has generated toponyms, such as *la Cendrosa* ‘land with the appearance of ash’, and the derivatives and compounds of *arena* (‘sand’; *Riudarenes*, etc.) or from the Latin *ARENEUM* (*Areny*, *Arenys*, *Balsareny*, etc.).

Another obvious group is that associated with water (*aigua*), which serves to designate not only hydrographic elements but also human settlements and other places, whether it be a reference to water in general terms (e.g. *la Bonaigua* ‘good water’, *Aiguaviva* ‘water rapids’, *Aiguatèbia* ‘lukewarm waters’), to stagnant waters (*la Llacuna*, *Llagunes*: from *llacuna* ‘pool, lagoon’) or to currents: *Riu* ‘river’, *Torrent* ‘intermittent river’ (cf. *la Riera*), etc. And we should not overlook springs, which are present in the toponymy both in their simple (*les Fonts* ‘springs’) and derivative forms

(through *fontana*: *Fontanet*, *Fontanals*, etc.) and in composition (*Fontllonga* ‘long spring’, *Fontdepu* ‘spring of the well’, etc.).

Similarly, both the local vegetation and fauna provide a large number of place names. Among the names of plants, along with countless references to trees, shrubs, fruits and plants, there are toponyms that allude to woodland in general; numerous examples of *selva* < lat. SILVA in the Catalan toponymy (including some compounds such as *Selvanera* < SILVA NIGRA ‘black, dense forest’) indicate the extent of the survival from mediaeval times of this type of lexicon compared to the standard *bosc*. As for references to more specific species, in addition to many single references (e.g. *l’Arboç* ‘strawberry tree’, *Folgueroles* < *falguera* ‘fern’), the duality between simple nouns (with generic value) and nouns derived by suffixation stands out (in particular the so-called *botanical groups*: cf. Bastardas 1994): *el Bruc* and *Bruguera* < *bruc* ‘heather’; *el Fonoll* and *Fonolleres* (and *Fonollosa*, etc.) < *fonoll* ‘fennel’; *Oliva* and *Olivella* < *oliva* ‘olive [olive tree]’; etc. The toponyms *Cercs*, *Cerqueda*, *Puigcercós* point to the survival in Romance of the Latin QUERCUS ‘oak’, while *Juncosa* and *la Jonquera* reflect the dialectal duality between the terms *junc* and *jonc* < lat. JUNCUM ‘rush’. Even marine vegetation is reflected in the toponymy: from *alga* ‘seaweed’ we get the name of *l’Alguer*, the easternmost point of the Catalan language domain, in Sardinia.

As far as the fauna are concerned, the references to common species in woods, meadows and mountains predominate: to the examples of suffixation (typically in *-era* < *-ARIA*) discussed in other sections, here we can add *Corbera* < *corb* ‘raven’, and the compound *Vall-llebrera*, from *vall* ‘valley’ and *llebre* ‘hare’. Masculine forms can be found (*Granollers* probably derives from *granolla*, an old dialectal variant of *granota* ‘frog’), and logically nouns with other suffixes (*Aguilar* < *àguila* ‘eagle’) or simply the noun without any derivation (*la Llagosta* ‘grasshopper’, *Cabra* ‘goat’, etc.).

2.2.2 Names derived from human activity

A similar degree of tangibility is found in the toponyms that refer to human activity, in particular names derived from buildings or other constructions. In some cases, this might be a specific element of the architecture: *els Arcs* and its diminutive *Arquells* refer to an *arc* ‘arch’, and *Mur* – like the Mozarabic *Muro* and others – is a direct reference to *mur* ‘wall’, and we also find *mur* (more specifically *muralla* ‘ramparts’) in *Montfalcó Murallat*, a fortified town. Other names, likewise, derive from *castell* ‘castle’ (among others we find the suffixed forms *Castelló*, *Castellet*, *Castellar*, and compound nouns, such as *Castellblanc*, *Castellfollit*, etc.). Similarly, other military constructions are referenced: *torres* ‘towers’, present as a simple noun in *Torres*, and in derivative and compound nouns (*Torroella*, *Torregrossa*, etc.). On the other hand, *palaus* ‘palaces’ refers to civilian architecture; yet, we would be wrong to picture grand royal palaces (although some are to be found, as in the hybrid Romance-Arabic compound noun, *Palau-solità* < PALATIUS SALATIN ‘palace of the kings’ [cf. *sultans*]), but rather

more modest buildings, as is suggested by the relative frequency of this root in rural areas, appearing both as a simple noun *el Palau*, and in compound forms (*Montpalau* ‘mansion hill’, *Palau-sator* < PALATIUS IPSA TURRE ‘tower’, etc.), and even names based on an ancient derivative *palol* < lat. PALATIOLUM (diminutive of PALATIUM ‘palace’) in such names as *Palol* and *Palou*.

Other names with links to human activity are those that refer to trade, industry, livestock, and agriculture, etc. Often, they include an allusion to a building or construction: thus, we find *hostals* ‘hostels’ (*els Hostals*, *l’Hostal Nou*), and the ancient *tenes* (or *tendes*: ‘shops’ or ‘stalls’) of the river *Tenes* and *Calldetenes* (a compound noun with *call* ‘neighbourhood’ or, more frequently, ‘Jewish quarter’). There are also clear references to commercial and manufacturing activities in *Malmercat* ‘bad market’, *Fornells* (diminutive of *forns* ‘kilns’), *Ferrerries* (‘smithies’), *Vidrerres* (furnaces for the production of *vidre*, ‘glassworks’), etc. It seems clear that honey was produced in *Arnes* in *arnes* ‘beehives’, as it was in *Piera* < lat. APIARIA ‘apiary’.

We also find names relating to a historical event or legal situation, whether it be a reference to the enjoyment of certain freedoms (the *Franqueses* or *Vilafranca* with allusions to *franca* ‘free’) or, just the opposite, alluding to some kind of feudal relationship; *Vila-real* is clearly the town of the king (*real* ‘royal’), and *el Marquesat*, *el Comtat*, *el Priorat* (the names of *comarques*) were, respectively, the domains of a *marquès* ‘marquis’, a *comte* ‘count’, and a *prior* ‘a prior of a monastery’. A confrontation over boundaries (*Contrast* ‘contrast, dispute’) was also a motive for designation, as was the fact that a bishop had his principal seat there (*la Seu* ‘cathedral’) or even a second residence or a property that he owned: *la Bisbal* ‘episcopal’; *Canòlic* < CANONICUM ‘of the canons’, etc.

Staying with religion, we find place names that also allude to religious buildings (e.g. *Monistrol* < MONISTERIOLUM < MONASTERIUM ‘monastery’) or other physical elements of religion (e.g. *Ares* – an *ara* being an altar often erected in a difficult mountain pass, where it was propitious to make an offering or some other ritual of gratitude or petition for divine protection). Another pre-Christian toponym is that of *Portvendres*, with its well-known etymology PORTU VENERIS ‘port of Venus’.

Typically, the toponymy of a religious nature presents a lower degree of tangibility than that which refers to the natural environment or to other human activities, insofar as the motif to which it alludes can be highly generic (*Montserrat* ‘holy mountain’, *Gràcia* ‘divine grace’, *Puiggraciós* ‘hill of grace’, etc.). This is evident in the case of hagiotoponyms, since it is very rare to find for example a *Sant Ramon*, where a historical figure known as Ramon became a saint and generated a toponym for his birthplace; on the contrary, most religious dedications respond to a decision which, if not entirely arbitrary (because there are always dedications that follow a fashion or a regional or chronological tendency), has at least little justification and is not especially objective. Proof of this is the repetitive, commonplace nature of certain names (*Sant Martí*, *Santa Coloma*, *Sant Joan*, *Sant Andreu*, etc.), and, for that reason, we comment below on just a few of the more interesting cases.

There are, besides the more obvious and clearer cases, some hidden hagiotoponyms, as well as those of a more obscure, enigmatic nature, and even some false hagiotoponyms. The hidden hagiotoponyms occur, for example, in those where the word *sant* (or its ancient variant *sent*) is agglutinated with the personal name or even with other complementary elements: *Santanyí* is, in reality, an alteration of *Santa Agnès*; *Sentmenat* hides the ancient *Sent Menat* (where *Menat* is a Greco-Latin anthroponymic derivative of *Menna*); *Santpedor* is the agglutination and contraction of *Sant Pere d'Or* (and, indeed, in that location there flows a *riu d'Or* 'river of gold', where in ancient times this metal was found). Presenting varying degrees of obscurity are the names of saints that have fallen into disuse, that is, those that are no longer used as given names (*Sant Cugat*), or which have undergone some sort of transformation, be it of the Catalan name (*Sant Guim* < *Guillem*; *Santa Càndia* < *Càndida*) or the Latin etymon (*Sant Oïsm* < *ONESIMUS*; *Sant Orenç* < *AURENTIUS*; *Santa Pellaia* < *PELAGIA* or *PALLADIA*), and even an agglutination with other elements, distorting the toponym to the point that the name of the saint becomes unrecognisable (*Sant Pesselaç* < *Sant Pere s'Arç*, that is, *de l'arç* 'hawthorn'). Finally, we find various false hagiotoponyms: this would appear to be the case for *Santa Linya* (if Coromines is correct in saying that there is no saint *Linya*, *Línia* or *Licina*, and that the name is, in fact, an alteration of *SALTA ILICINEA* 'wood of holm oaks') and perhaps also for *Sant Domí* (if it does not allude to a saint *Domici* < *DOMITIUS*, but rather hides the Germanic anthroponym of *Sendomir*, derived from the base term *SINDEMIRO*).

2.2.3 Place names derived from personal names

In examining the names of saints, we have entered once more into the field of personal names, and here we see that both types of name present the same relative degree of arbitrariness (or objectivity). That is, the process by which an anthroponym becomes a toponym – usually a reference to the founder or former owner of a place – retains a link with reality, but clearly this association is not as strong as might be the case for descriptive names referring to the landscape or the surrounding vegetation, which are unlikely to experience significant changes over the centuries. Up to this juncture, we have examined many toponyms derived from anthroponyms; many of the Latin and Arab names fall in this category as do, without exception, all Germanic names. Here, therefore, we shall limit ourselves to mentioning just a few more. For example, those names that, in principle, are not attributable to any of these other source languages, but which appear to be derived directly from Catalan (such as the frequently occurring *Llorenç*) or from French (the *Lluís* in *Montlluís*, in French Catalonia, responds to a military fortification built on the orders and in honour of King Louis XIV), and even from Hebrew (*Sabadell* < *SABBATELLUS*, perhaps the diminutive of an anthroponym *Sàbat* < *SABBAT* 'Saturday').

If we return to the names of Latin origin, it is worth adding at least a couple of derivatives ending in -UCIUS (*Urús* < AURUCIUS; *Bellús* < BELLUCIUS) and, as a special case – but one that is relatively frequent in Valencia – derivatives ending in -ANUM which, instead of presenting the typical ending -à, evolved, due to the influence of Arabic (or perhaps Mozarabic) to form the ending -ent. This explains the well-known toponyms of *Crevillent* < CALVINIANUM, *Ontinyent* < ANTONIANUM, *Lutxent* < LUCIANUM, etc.

Halfway between Roman and Germanic anthroponymy (in fact, originating from the Celtic world, probably from Gaul), we find the suffix -ac < -ACUM, which has on occasions been considered a variant of the above, given its parallel behaviour (cf. Aebischer ¹1926). It is found in a few place names that are also undoubtedly anthroponyms, and, significantly, of diverse origins. They include *Reixac* < RIXIACUM, *Franciac* < FRANCIACUM, and *Vulpellac* < VULPILIACUM.

Finally, place names that are anthroponyms and of clear Germanic (Gothic or Frankish) origin, apart from those mentioned in the corresponding section, include, among many, those of *Manlleu* < MANILEUB; *Mataró* < MATHERION; (*la*) *Geltrú* < WISALTRUD; *Sunyer* < SUNJAR. Another obvious example is *Renau*, which, despite its resemblance to *Renan*, does not derive from the same base of RAGINAND, but rather from the parallel base of RAGINALD.

2.2.4 Place names based on other place names

A lower level of objectivity – but not to the extent of extreme arbitrariness – is exhibited by place names that are derived from other place names. We have already seen this phenomenon in section 2.1.6 above, where the examples given corresponded to place names originating from adstrate languages. This was the case, for instance, of toponyms originating from Occitan (*Verdú*, etc.), but in reality – and for similar causes, during the territorial expansion of the 12th and 13th centuries – the phenomenon is also found within the same territories of the Catalan language. Thus, we can speak of ‘duplicate’ names, some of which concentrate around the capital of *Catalunya Nova*, Lleida (*Bell-lloc*, *Montcada*, *Rosselló*, *Puigverd*, *Artesa de Lleida*, replicas that is of place names lying further to the northeast of the territory), while others are to be found in Valencia (another *Montcada*, *Montserrat*, the *Girona* river, etc.). More unusual cases include *Boldú*, which derives from the ancient form of *Besalú* (*Besaldú*), and *Novelda*, which incorporates an initial adjectival element, *nova* ‘new’, in front of the not too distant name – also Valencian – of *Elda*.

Likewise, in Catalonia, frequent use is made of the diminutive suffix to indicate this duplication, generating what we designate as *diminutive toponyms*. Some correspond to adjacent neighbourhoods (*Avinyonet*, *Pierola*, *Lloberola*) or even to a quarter within that place itself (the case of *la Barceloneta* in Barcelona). In other cases, however, because of the distance between the referents, the diminutive value is quite

weak (*Falset*, *Gironella*, *Cerdanyola* are, in spite of everything, diminutives of *Fals*, *Girona*, *Cerdanya*, respectively). We also know, thanks to ancient documentation, that some diminutive forms (*Vilagrasseta*, *els Omellons*, *Agramuntell*, today *Gramuntell*) were not originally known as such, but rather incorporated this suffix some decades, or even centuries, later.

Finally, some toponyms are formed from the demonyms associated with other place names – be it from distant (*Gallecs* ‘people of Galicia’; *Navars* ‘people of Navarre’) or quite close references (*els Pallaresos* ‘people of Pallars’).

Whatever the case, toponyms of this kind retain a link to their referent or, at least, contain a record of their origins, a non-random source. The maximum degree of arbitrariness occurs in the so-called case of the *toponymic series* (to use the terminology proposed by Moreu-Rey ²1982, 160), evident in urban nomenclature (that is street names that, for no apparent reason, refer to the same semantic field), but there is a debate as to whether this procedure may not also have occurred in the major toponymy, especially in a context of extensive territorial repopulation. For example, the similarities shared by the names of the Cistercian foundations – *Vallbona*, *Vallsanta*, *Vallclara*, *Valldigna*, etc. – would seem to correspond clearly to just such a *toponymic series*.

2.3 Morpholexical typology of Catalan toponyms

Finally, Catalan toponyms – in common with all other toponyms – can also be analysed in terms of their morphology, that is, according to the type of lexical formation they present. Three types can be distinguished: simple, derivative and compound nouns. This classification, as we have illustrated already with numerous examples, is independent of the two outlined in the preceding sections but can be interlaced systematically, so that each place name presents or should be able to present a triple typology: historical (that is, in relation to the linguistic stratum in which it was created), semantic (with regards to the referent to which it alludes) and morpholexical (with regards to the formal procedure it employs).

2.3.1 Simple nouns: substantives, adjectives and others

Toponyms formed from a simple lexical base or derived by suffixation are not unusual. Among these the most typical are those that employ a common substantive (by general rule, geographical in nature); i.e. *la Roca* ‘rock’, *Horta* ‘area of horticulture’, *Ponts* ‘bridges’, *Parcs* ‘walls’, *Sitges* ‘silos’, etc.

Less frequent is the use, as simple lexemes, of adjectival terms (e.g. *l’Aguda* ‘sharp’, *els Aspres* ‘rough’) or participles (i.e. *el Tallat*, *la Tallada* ‘cut’ *masc.* and *fem.* forms, *Foradada* ‘tunnelled’), in which we need to see the underlying presence, elided

in a prototoponymic phase, of a generic noun to which the adjective was applied: 'la [mountain] aguda', 'la [rock] foradada', 'els [lands] aspres', etc. This was already occurring in Latin: *Caldes* < lat. CALIDAS 'warm', undoubtedly corresponding to the [ACQUAS] CALIDAS 'warm [waters]'. Indeed, the various towns (no less than four) known as *Caldes* are traditionally associated with thermal baths.

2.3.2 Derivative nouns: suffixation (and prefixation) in Catalan toponymy

Many place names have been created by derivation, above all by suffixation. Prefixation, in contrast, is unusual in toponymy, although not completely non-existent: e.g. *Sopeira* < SUB PETRAM 'under the rock', or *Tresserra* < TRANS SERRAM 'behind the mountain range', in addition to a series of Pyrenean names beginning in *En-* < lat. IN (e.g. *Encamp*, *Engordany*), which Coromines (1970, 9–11), and a few other scholars, studied.

Derived from a lexical base that might, for example, have a botanical reference (cf. Bastardas 1994), or that of another semantic class, we typically find toponyms that make up authentic suffixal series. This is the case of derivatives ending in *-et* and *-eda* (< lat. -ETUM/-ETAM), with a collective value, which should not be confused – especially in the masculine form, given their evident homonymy – with those that contain the common diminutive form *-et* (feminine *-eta*); e.g. *Lloret* < *llor* 'laurel', *Maçanet* < *maçana* 'apple' (or, in the feminine form, though here less confusing, *Pineda* < *pi* 'pine', *Fenolleda* < *fenoll* 'fennel', etc.). A similar case is provided by that of the suffix *-era* (< lat. -ARIA), which is usually applied to zoonymic lexemes (e.g. *Cabrera* < *cabra* 'goat', *Llobera* < *llop* 'wolf', etc.), as well as to lexemes from other semantic fields (e.g. *Porrera* < *porro* 'leek', *Peguera* < *pega* 'pitch'). It has a similar value to the suffix *-ós* (< lat. -OSSUM), which is more frequent in the feminine form and usually applied to plant names (e.g. *la Molsosa* < *molsa* 'moss', *la Cardosa* < *card* 'thistle') but also to the land (e.g. *la Cendrosa* < *cendra* 'ash') and even to meteorological phenomena (e.g. *les Ventoses* < *vent* 'wind'). In fact, these formations are adjectival in form, making it necessary to suppose that a noun of a generic nature has been elided: 'les [terres 'lands'] ventoses' (cf. the compound noun *Puigventós* 'windy hill'), 'la [torre 'tower'] cardosa' (with thistles), etc.

2.3.3 Compound nouns: noun compositions in Catalan toponymy

Finally, noun composition is also a typical procedure in the formation of Catalan toponyms, both between names created directly in Catalan and between those that come from other strata (e.g. the Latin place names of *Igualada* and *Morvedre*, the Arabic names of *Sidamon* and *Massalcoreig*, etc.). We also find compound place names formed from a generic substantive element (of a clear Romance origin) and an anthro-

ponym, which may well not be Latin (as it is in *Palafrugell* < PALATIUM FRUGELLI, and *Vilobí* < VILLA ALBINI) but Arabic (*Castigaleu* < CASTRU HALEF) or, often, Germanic, as in the case of the toponyms *Montmaneu*, *Viladot*, *Castelladral*, *Torreneral*, etc., in which we find the Germanic anthroponyms of *MENED*, *OTO*, and *EDRAD*.

One of the recurring models, among others, is that which combines a verbal component and a substantive form (*Miralcamp* ‘look over the field’, *Guardamar* ‘watch over the sea’), or an adjective with a noun (*Alta-riba*, *Setcases*, and the series based on the adjective *bell* ‘beautiful’: *Bellpuig*, *Bellmunt*, *Bellprat*, etc.). More frequent, however, are compound nouns that present these elements in the inverse order (i.e. substantive noun and adjective): *Montsec*, *Puiggròs*, *Vallcorba*, *Torreblanca*, *Castellnou*, etc. Compound nouns are also formed between two substantive elements (*Matadepera* ‘wood of stones’, *Vilademuls* ‘town of mules’), typically without a prepositional nexus (*Montseny*, *Pedraforca*, *Peramola*).

Finally, mention should be made of non-agglutinated compound nouns, that is, toponyms that, for traditional (or pragmatic) reasons, are not written as one word: *Molins de Rei*, *la Vila Joiosa*, etc. (and which in certain contexts are often referred to simply as “Molins” or “la Vila”). But we should distinguish these names from those which, because of the occurrence of place name homonyms, present a complement with little use outside that of officialdom for purposes of clarification of reference: *Puigverd d’Agramunt* and *Puigverd de Lleida*, *la Bisbal d’Empordà* and *la Bisbal del Penedès*, etc.

3 Catalan anthroponymy

In the same way that the toponymy of a territory is a reflection of its history and geography, its anthroponymy can also be understood as a mirror of both the past and present of its culture and society, channelled through the territory’s own language or through other languages that have been used there or that have a presence there. Anthroponymic systems have been subject to constant evolution, making their history extremely complex. Marked swings in the pendulum have occurred from the early use of single, unique, primitive names (e.g. the Iberian name *Indibil*) to the Roman adoption of complex names (e.g. *Caius Vibius Lupercus*), again from the unique Visigothic names (e.g. *Baddo*) to complex Arabic names (e.g. *Abu Bakr al-Balagî*), and then back to the single names employed by the early Catalans (most still in the Visigothic tradition: e.g. *Sunifred*) and the recovery, between the 11th and 13th centuries, of complex names, with a forename and surname (or two surnames, in more modern times) and the possibility of an additional nickname.

Thus, the system currently in use in the Catalan-speaking territories – with some legal distinctions depending on the modern nation to which each region belongs – is the result of an evolution that, without our needing to trace its history back to the Iberians, Romans or Arabs, presents a well-defined historical continuity dating back

to the 9th–10th centuries, the period coinciding with the creation of the Catalan counties and the beginning of its medieval territorial expansion. The system evolved from one characterised by the use of a unique personal name, with a predominance of Germanic names (Visigothic influenced by a pre-Arabic or Frankish tradition, given the region's initial dependence on the Carolingian Empire), together with some names of Latin or Christian tradition (i.e. biblical or Greek names), as is highlighted by the abundance of examples cited in the *Repertori d'antropònims catalans (RAC-I)*, which includes all the names documented until the year 1000. This primitive system gradually acquired greater complexity throughout the 11th century, when some second names appeared (a kind of nickname, indicating an individual's trade or parentage), a practice that became more widespread in the 12th century (cf. Aebischer 1928, a study whose importance remains undiminished 90 years after its initial publication). It was not until the 13th century, however, that the definitive transformation was made to names becoming hereditary (that is, transmitted from parents to children, regardless of their semantic content), and the birth of surnames as we know them today occurred (see Moll ²1982).

3.1 Typology of Catalan forenames

When analysing forenames, one cannot help but notice that, unlike toponyms and other anthroponyms, few Catalan given names actually originated in the Catalan language. This reflects the specific characteristics of forenames, including the fact that naming is performed on an individual basis and that this process is strongly influenced by shifting fashions from one generation to the next often in response to social trends. Names, moreover, represent strong local and regional traditions, most notably the dominance of the Catholic saints over centuries, and the persistence – even after centuries, for the reasons cited above and, no doubt, for other motives – of names created in languages that have disappeared or which are no longer used in the territory. The result is a system that, rather than presenting obvious strata, responds to a series of cultural or religious influences.

For centuries, the most common forenames among the Catalan population have been – and in part remain – those of biblical tradition (*Josep, Joan, Jaume, Joaquim, Miquel; Maria, Anna, Marta*); Greek, be they biblical or Christian (*Jordi, Esteve, Andreu; Eulàlia, Bàrbara*); and more rarely those that are classical or mythological in origin (*Alexandre, Narcís, Jacint; Helena, Eugènia, Ariadna*). Likewise, we find Latin, or Roman, names (*Antoni, Marc, Víctor; Júlia, Martina*), Christian names (*Pau, Pere, Magí, Agustí; Cristina, Beatriu*), and Germanic names (*Lluís, Carles, Ramon; Berta, Elisenda*).

Names with other origins are also to be found, including Italian (*Francesc, Gemma*), Occitan (*Elionor, Mireia*), French (*Eloi, Joel*) and, of course, Spanish (*Isidre, Diego, Domingo, Teresa*). A few years ago, names from the north of the continent

became fashionable, including those of Slavic (*Olga, Ivan*), Scandinavian (*Eric, Axel*), and also Basque origin (as well as the popular *Xavier, Aitor, Unai, Arantxa* and *Itziar* are all gaining in popularity).

Alongside these, the purely Catalan names are largely limited to concepts of a more mystical nature (*Salvador* ‘saviour’, *Concepció* ‘conception’) and to certain advocations of the Virgin Mary, be they abstract (*Dolors* ‘sorrow’, *Esperança* ‘hope’) or substantiated via some specific element (*Neus* ‘snows’, *Roser* ‘rosebush’). Only a few basic terms (*Blanca* ‘white’, *Marina* ‘marine’, *Rosa* ‘rose’) can be added to the list of etyma that are readily understood by everyone.

3.2 Typology of Catalan surnames

Catalan surnames exhibit a typology that is common to that of the rest of the Romance-speaking and Western worlds, with five main groups: patronyms; names indicating a trade; descriptions of a personal trait; allusions to territorial origin; and references to a place in the same settlement (or district). In all cases, these motifs were used as distinct, defining symbols of a particular individual in relation to his or her fellow inhabitants when surnames first appeared (initially, in the form of nicknames).

First, we find the patronyms, that is, surnames based on the name of the father, either by attaching a suffix (as is typical in Spanish with the atonic termination *-ez*) or, as was the norm in Catalan, without any formal variation. Thus, many first names still common in Catalan today have, by this same channel, come to us as surnames: *Agustí, Andreu, Bertran, Martí, Rafel, Salvador, Tomàs*, etc. Of course, this phenomenon is completely independent of the etymological origin of these names, so we find names that are Latin (*Marc*), biblical (*Gabriel*) and Germanic (*Lluís*) in origin. This group also contains a number of ancient anthroponyms, which were common in mediaeval times, but that subsequently became infrequent or obsolete as forenames. Some have a Latin etymon (e.g. *Nadal, Vidal, Fortuny*) but most are Germanic etyma given the preponderance of these anthroponyms in the Middle Ages (e.g. *Aimeric, Gibert, Gomà, Guiter, Mir, Riquer*).

The second, and no less typical, group is that of surnames referring to trades. Many are obvious to the extent that they allude to trades still very much practiced, such as *Ferrer* ‘smith’, *Forner* ‘baker’, and *Fuster* ‘carpenter’ (also, of course, without the suffix *-er*: *Mestre* ‘master’, *Pagès* ‘farmer’, *Sastre* ‘tailor’, etc.), but others refer to ancient occupations that have long since disappeared, such as *Bover* ‘cowherd’, *Cabrer* ‘goatherd’, and *Traginer* ‘carter’.

The third group is made up of surnames that referred to a personal – usually physical, but sometimes moral or behavioural – trait (i.e. *Alegre* ‘merry’, *Dolç* ‘sweet’, *Moll* ‘soft’). Alternatively, the name might have been a reference to the colour of the person’s skin or hair (*Blanc* ‘white’, *Bru* ‘dark’, *Ros* ‘fair’, *Roig* ‘red’), their body shape (*Fort* ‘strong’, *Gras* ‘fat’, *Prim* ‘thin’) or to some other notable feature (*Esquerrer* ‘left-

handed', *Morro* 'snout', meaning 'prominent lips', *Tort* 'crooked'). This group might also include surnames based on animal names (e.g. *Bou* 'bullock', *Colom* 'pigeon'), the individual having borne some resemblance to the animal in question.

A fourth group includes the surnames that refer to the origins of the individual or his or her family, either by using a national (*Alemaný* 'German', *Francès* 'French') or regional demonym (*Cerdà* 'from Cerdanya', *Pallarès* 'from Pallars'), or, more directly, the name of the country (*Aragó* 'Aragon', *Castella* 'Castile'), the *comarca* (*Garrigues*, *Segarra*, *Vallès*) or even a river (*Cinca*, *Giüell*, *Llobregat*). More frequent, however, are references to villages, in some cases via the corresponding demonym (*Berguedà* 'from Berga', *Tolsà* or *Toldrà* 'from Toulouse'), but primarily through direct reproduction of the name of the settlement, as in the case of *Alcover*, *Balaguer*, *Camarasa*, *Cardona*, *Ivars*, *Ripoll*, and many more. Places outside the linguistic domain are also recorded in Catalan surnames. This is the case for the Aragonese towns of *Terol* (Teruel) and *Montsó* (Monzón) and the Occitan towns of *Pàmies* and *Foix*.

Finally, in the fifth group, we find surnames that refer to places or elements of the same village or in its surrounding neighbourhood or district, including, for example, *Puig* 'hill', *Serra* 'mountains', *Vall* 'valley', *Pla* 'plain', *Riu* 'river', *Font* 'spring', *Vila* 'town', *Torre* 'tower', *Mas* 'farmhouse', *Pou* 'well', *Porta* 'door', *Pont* 'bridge', etc. In fact, these and many others that can be similarly categorised (often also in the plural: *Pous*, *Rius*, *Torres*, etc.) are the most frequent names of Catalan origin.

The above categories, however, do not exhaust all surname types. For example, we find some that refer to a familial relationship (*Nebot* 'nephew', *Hereu* 'first-born') or to a social position, be it real or metaphorical (*Rei* 'king', *Duc* 'duke'), and, above all, those that designate some particular circumstance: e.g. *Abril* 'born in the month of April', *Salvat* 'rescued/saved', etc.

Naturally these classifications can be questioned, especially as many surnames could be interpreted in more than one way. For example, *Segur* 'trustworthy' might refer to a personal trait or to any one of the settlements with that name. Likewise, *Foix* is the name of an Occitan town, but also of a Catalan village and river, etc. There are clearly many variables and many ways of classifying them. After all, we are not dealing with an exact science; on the contrary, we find ourselves immersed in one of the most human, and personal, of the human sciences.

4 Bibliography

- Aebischer, Paul (2006, 1926), *Études de toponymie catalane*, Barcelona, Institut d'Estudis Catalans.
- Aebischer, Paul (1928), *Essai sur l'onomastique catalane du IXe au XIIe siècle*, Barcelona, Biblioteca Balmes.
- Balari Jovany, José (2016, 1899), *Orígenes històrics de Catalunya*, Sant Cugat del Vallès, Instituto Internacional de Cultura Románica.
- Barceló, Carme (2010), *Noms àrabs de lloc*, Alzira, Bromera.

- Bastardas, Maria-Reina (1994), *La formació dels col·lectius botànics en la toponímia catalana*, Barcelona, Reial Acadèmia de Bones Lletres de Barcelona.
- Coromines, Joan (1965–1970), *Estudis de toponímia catalana*, 2 vol., Barcelona, Barcino.
- DCVB = Alcover, Antoni M./Moll, Francesc de B. (1926–1968), *Diccionari català – valencià – balear*, 10 vol., Palma, Moll.
- Förstemann, Ernst (1900), *Altdeutsches Namenbuch: Personennamen*, Bonn, Hanstein.
- Moll, Francesc de B. (²1982, ¹1959), *Els llinatges catalans*, Palma, Moll.
- Moreu-Rey, Enric (1974), *Toponímia urbana i onomàstica vària*, Palma, Moll.
- Moreu-Rey, Enric (1981), *Renoms, motius, malnoms i noms de casa*, Barcelona, Millà.
- Moreu-Rey, Enric (²1982), *Els nostres noms de lloc*, Palma, Moll (¹1965 as *Els noms de lloc*).
- Moreu-Rey, Enric (1991), *Antroponímia. Història dels nostres prenomes, cognoms i renoms*, Barcelona, Universitat de Barcelona.
- OnCat = Coromines, Joan (1989–1997), *Onomasticon Cataloniae. Els noms de lloc i noms de persona de totes les terres de llengua catalana*, 8 vol., Barcelona, Curial.
- RAC-I = Bolòs, Jordi/Moran, Josep (1994), *Repertori d'antropònims Catalans*, vol. 1, Barcelona, Institut d'Estudis Catalans.
- Velaza, Javier, et al. (2002), *Els substrats de la llengua catalana: una visió actual*, Barcelona, Societat Catalana de Llengua i Literatura.

16 Translation

Abstract: The history of translation, like the history of language and literature, is interrelated with the history of the country in which translations are being produced, and in the case of Catalan-speaking territories, this relationship is an extremely dynamic one. The history of these regions explains the profusion of translations from the Middle Ages and their in some cases pioneering character, the drop-off in the number of translations from the 16th to well into the 19th century (outside Northern Catalonia and Menorca), the revival during the first few decades of the 20th century followed by further decline during the Franco regime and the stabilisation that is in evidence today.

The translation of Catalan literature into other languages has been largely dependent on the prestige it has enjoyed during each period and is closely linked to the rise of the Catalan language itself.

Keywords: translation into Catalan, translation from Catalan, medieval translation, modern translation, contemporary translation

1 Introduction

As in the case of many other languages, particularly during their early development, translation has played a decisive role in the establishment and development of the Catalan language. The Toledo School of Translators, for example, was highly significant in the shaping of Spanish, just as translations of the classics influenced La Pléiade in their model of French.

Catalan experienced its golden age during the Middle Ages when it was the language of the court and was used for all kinds of legal and public documents; it was the language of the Royal Chancery, the governments of Catalonia, Valencia and the Balearic Islands, and during this period also spread to Italy and Greece. However, as a result of Spanish hegemony during the 16th and 17th centuries, along with laws restricting the use of Catalan in the 18th century and during the Franco dictatorship in the 20th century, the use of Catalan as a language of culture later declined.

However, the rich medieval tradition and the Catalan *Renaixença* in the 19th century managed to keep the language alive over the centuries. Catalan was the first language of the Catalan people, who were largely monolingual practically up to the beginning of the 20th century. At this time, political and social advances called for the establishment of “standard Catalan”, which was provided by Pompeu Fabra. In a situation of marked diglossia, translation became a tool for regularisation and a

<https://doi.org/10.1515/9783110450408-024>

means of assembling a language model that was fluent and efficient, and which could be used in all areas of life. The translation of masterpieces and new foreign works into Catalan was a way of strengthening the target language and enabling the subtleties of art and thought to be expressed through it. Above all, it was a way of restoring Catalan's status as a language of culture. The history of Catalan translation has always taken this into consideration, as will be discussed in this article.

2 The Middle Ages

As was the case for a number of other European languages, the first documents to be written in Catalan in the 12th and 13th centuries were translations: law, science and religion provided a sense of belonging to the community expressing itself in Catalan. The classics had already begun to be translated well before the *Renaixença*, and from the mid-14th century a group of translators made up of administrators working for the Royal Chancery were able to use a fairly standardised form of Catalan as a target language thanks to the official nature of their work. The court was interested in the classics of Ancient Greece and Rome and in 1383 Titus Livius, mentioned as early as 1315, was translated, as were Palladius, Cicero, Valerius Maximus, Ovid and Seneca a little later. In fact, in quantitative terms, '78 per cent of Catalan medieval translations are from Latin' (Pujol 2002, 11). During this period Latin was considered *superior* to the *vulgar* Romance languages that had recently emerged and which had neither tradition nor authority.

The Italian connection led to a Catalan translation of Petrarch (the humanist and moralist who wrote in Latin, not the Tuscan poet and writer of the *Canzoniere*) a few years after his death. In 1388, Bernat Metge translated *Griseldis*, the last book of Boccaccio's *Decameron*, not from the original but from Petrarch's Latin translation. It was the first translation of a work by Petrarch produced on the Iberian Peninsula. Sometime later, the friar Antoni Canals produced a prose translation of the *Parlament de Aníbal e Scipió* from Petrarch's *Africa*, and Catalan became the first language to rewrite verse in the ultimate metrical form of the original with the translation of Dante's *Commedia* by Andreu Febrer, poet and official in the court of King Alfons the Magnanimous. In the same year the monks of Sant Cugat monastery translated Boccaccio's *Decameron* in its entirety. An unknown translator also rendered Boccaccio's *Fiammetta* while another translator, supposedly Narcís Franch, produced a Catalan version of *Corbaccio*.

Throughout the 15th century, Latin gradually ceased to be a widely read and understood language, which led to the need for a new method of translation, and one which did not follow the original text as literally. In other words, the perception of Catalan as a self-contained language began to emerge, with all its quirks and challenges compared to Latin, which never ceased to be considered the most perfected language of all. Translators recounted this gradual recognition of the *vulgar*

language in several prologues. The best examples are those by Ferran Valentí in his rendering of the *Paradoxa Stoicorum* by Cicero (around 1450), Arnau d'Alfarràs in the *Regula Benedicti* by Saint Benedict (1457) and Francesc Alegre in the *Metamorphoseon* by Ovid (1494) (Badia 1991).

In this brief overview of medieval translations, one in particular should be singled out. In 1478 a complete edition of the Bible was printed in Valencia, of which only the last page, containing the colophon, remains. This is most likely the result of the prohibition of biblical texts in Romance languages from the 13th century onwards and, in particular, the establishment of the Inquisition. This makes Catalan the fourth language to be used for a complete translation of the Bible after German, Italian and Dutch. Even though Bonifaci Ferrer is credited with translating it from the Latin *Vulgata*, it would seem to have come from a pre-existing version worked on by a group of translators of Jewish origin.

3 The Modern Period

It is well known that the 16th and 17th centuries have been referred to as the *decadència* in terms of the Catalan literary scene. In fact, while production diminished quantitatively, it was a decline in the ambition of the works being produced as a result of Catalan's lack of prestige as a language that was most noticeable. Catalan was increasingly undervalued in the face of the elevated use of Castilian, the language of the court. Given this context, it is hardly surprising that the need for translations into this underrated language all but disappeared, with the exception of popular texts and manual work, where it was still needed to be able to reach a broad public. So, religious works, traditional narratives and books for teaching and training continued to be translated.

Throughout the 18th century new social, cultural and political conditions allowed for different translation patterns in the two Catalan territories that were not at that time under the control of the Spanish crown: Northern Catalonia and the Island of Minorca.

As a result of the Treaty of the Pyrenees (1659), Catalonia had been divided into two states. On the French side, over a century later, most of the rural and largely illiterate population still only spoke Catalan. This led to a flourishing number of translations of classical works of French 17th century theatre. The main works appeared in what was known as the Tuïr group, led by Guillen Agel in the second half of the 18th century (Vila 1996).

For almost all of that century, Minorca was a colony of the British Empire and experienced significant development, its literature becoming imbued with highbrow European neoclassical trends. This led to a growing number of translations of theatre, which provided the best source of entertainment for a prosperous, restless and essentially isolated society. Educated in this atmosphere of reform, Antoni Febrer i

Cardona and Vicenç Albertí produced all of their work in the first decades of the 19th century, just a few years after the island passed back into the hands of the Spanish crown. Febrer i Cardona made his priority the translation of the Greek and Latin classics: Cicero's philosophical treatises, Phaedrus's fables and Virgil's eclogues. Between 1815 and 1820 Albertí translated around thirty works for the theatre (although only fifteen have been preserved) by Goldoni, Molière, Metastasio, Beaumarchais and Moratín (Paredes Baulida 1999).

4 *Renaixença?*

During the first few decades of the 19th century the level of confidence in Catalan as a literary language continued to decline. This low prestige had a knock-on effect on translations and there were very few produced until the final decade of the century.

One, however, stands out particularly for its scope. Josep Melcior Prat, who was exiled to England, translated *Lo Nou Testament*, which was printed in London in 1832 (one year before Aribau's famous work *La Pàtria*, which was considered to mark the beginning of the *Renaixença*; 713 *Renaixença*). Four editions were printed with a total of 20,000 copies (a considerable number, bearing in mind the percentage of the population that was literate).

A couple of years later, in 1834, one of only a handful of non-religious books to be translated during almost the entire 19th century appeared: *La noia fugitiva*, a version of *La Fuggitiva* by Tommaso Grossi. The translator was named as Joan Cortada, but it seems more likely to have been the work of Miquel Anton Martí, who also produced other unpublished translations of the Italian poets.

From the middle of the century, as a result of personal relations established between some of the Occitan and Catalan writers in a common climate of literary awakening in both regions, and the influence of the work of Mistral, works from writers on both sides were translated in and out of both languages. Just two years after the appearance of the original, Francesc Pelai Briz began to work on his Catalan version of *Mirèio* (1861) in verse, accompanied by his Spanish interpretation in prose (a good indication of the degree of diglossia that Catalan had achieved).

5 Up to the Civil War

The political and linguistic awareness that spread across Europe with Romanticism, and the central idea of the mother tongue as the national language, did not reach Catalan territory until much later, at the end of the 19th century during the belle époque of Catalan *Modernisme*. Despite their late arrival, these ideas had an unusually strong impact. The increase in literary production by such brilliant writers as the playwright Àngel Guimerà, the novelist Narcís Oller and the poet Jacint Verda-

guer, supported the conviction that the Catalan language had to be recovered and that it would be capable of transporting and broadening the local culture through the inclusion of some of the greatest literature of all time.

The magazine *L'Avenç* became the intellectual nucleus of this movement, publishing seven magazines along with 525 books in twelve collections. The Biblioteca Popular de L'Avenç was the longest-running collection, lasting from 1903 to 1915, and would also publish the most titles (147) and the most translations (55). The translated works reflected an eagerness to include all outstanding texts from different countries and periods: Dante, Shakespeare, Pascal, Goethe, Molière, Goldoni, Leopardi, Novalis, Tolstoy, Whitman and Ibsen, among others. The translators included some of the most important figures, both young and old, in the Catalan arts scene: Narcís Oller, Joan Maragall, Joan Puig i Ferreter, Pompeu Fabra and Joaquim Ruyra (Pla i Arxé 1975).

With the aim of creating a national theatre, Adrià Gual founded the Teatre Íntim, a Catalan version of the Parisian Théâtre d'Art, in 1898. The theatre staged plays by a number of foreign playwrights such as Aeschylus, Sophocles, Shakespeare, Molière, Goethe, Ibsen and Beaumarchais, with Gual translating some of the plays from French himself.

The desire to *catalanise* great works of literature can be seen most clearly in the approach to Shakespeare. Some forty translators have tackled his works over more than a century, from the first attempts to translate fragments of *Hamlet* in 1880 to the complete works by Salvador Oliva (1984–1992) and ten new translations by Joan Sellent (2016) and five by Miquel DescLOT (2017). There have been plenty of other projects in-between: from 1907 to 1910 the Biblioteca Popular dels Grans Mestres published sixteen Shakespeare plays, translated by different translators who often used the French versions as source texts; the poet Magí Morera i Galícia translated six of them directly from English from 1912 to 1924; before the civil war, Cèsar-August Jordana completed twelve for a fateful project that was later cut short; and after the war, Josep M. de Sagarra translated twenty-eight, which were later published and performed frequently up until a few years ago.

This considerable corpus includes everything from the erudite translations by Cebrià Montoliu of *Macbeth* (1907) and *Anfós Par* of *King Lear* (1912), for example, to the translations made more deliberately for the stage, such as those by Sagarra and Sellent. *Hamlet* has been the object of at least eight translations, while many of Shakespeare's other works have also been translated more than once, in particular *Macbeth*, *Othello* and *King Lear*. The sonnets have also aroused considerable interest from translators offering different interpretations: Ramon Font i Preses translated a selection for the magazine *Catalunya* in 1903; Morera i Galícia translated another sample in 1912 and 1913; in 1928 Carme Monturiol became the first to re-write the complete series; in 1958 Joan Triadú translated forty; Gerard Vergés completed another version of the whole series in 1993 as did Salvador Oliva in 2002, Txema Martínez in 2010 and, finally Salvador D. Insa produced a new version in 2016 (Buffery 2007; Pujol 2007).

Another case also corroborates the argument that translation not only satisfied a peremptory need to understand the foreign works, but also to win the language greater prestige. Over a period of 48 years at the end of the 19th and beginning of the 20th centuries, three complete Catalan translations of Cervantes' famous work *Don Quijote de la Mancha* appeared, while two more complete new versions were produced in 1969 and 2005. The total figure, including partial translations and adaptations, is close to fifty (Bacardí/Estany 2006).

Translation also crossed literary genres and artistic boundaries. Barcelona's passion for Wagner was partly ideological; it represented an alternative model to the *flamencoism* of Spanish clichés and became an assertion of Catalanism. In 1901 the Wagnerian Association was established and focused intensively on the dissemination and translation of the composer's works. Joaquim Pena, Antoni Ribera, Xavier Viura and Jeroni Zanné were commissioned to translate all of Wagner's works and adapt them to music, sometimes working with other collaborators (Janés i Nadal 1983).

As mentioned above, aesthetic and theoretical interest in translation really began with Catalan *Modernisme*. This was followed by *Noucentisme*, which established the basis for a true cultural policy. These movements created intellectual networks that managed to combine great philological rigour with the development of a cohesive group, and which were only broken up by the outbreak of the Spanish Civil War and the subsequent dictatorship (Murgades 1994).

The archetypal works of the *Noucentisme* period have to be those of Carles Riba. Following in the footsteps of Antoni Bergnes de les Cases in the 19th century, Riba incorporated the Hellenistic tradition into Catalan and in doing so created a style which went on to inspire other writers and translators such as Josep M. Llovera and Manuel Balasch, culminating in the adaptation of classical rhythms. A brilliant translation theorist, Riba understood the practice of translation as reading and interpretation, whilst also distinguishing between 'erudite' or scholarly translation and 'substitutive' or communicative translation. He worked in several languages and genres, translating the *Song of Songs* and the *Book of Ruth*, the great Greek tragedies in poetical and philological versions, Homer's *Odyssey* in two versions which have become part of the canon, the entire *Parallel Lives* by Plutarch, stories by Poe, the poetry of Hölderlin, Rilke and Kavafis, among others. It should be added that he also promoted and corrected many other translations, proposed and drew up directives for effective cultural policy in the field of translation and even revised versions of his own works in other languages (Malé 2006).

Translation of the Latin and Greek classics crystallised in the collection of the Bernat Metge Foundation, which took its inspiration from the Collection Budé, the Loeb Classical Library and the Teubner, and, *avant la lettre*, trained a veritable school of translators directed by Carles Riba. Since its establishment in 1928, the foundation has been responsible for the publication of more than four hundred volumes, except during the period 1939–1946, making the Greek and Roman classics highly accessible

to Catalans and, when there was no parallel series available, to Spanish readers (Franquesa 2013).

The leading modern writers in Europe were also translated into Catalan. Engaged in this endeavour were a number of active publishing houses such as Editorial Catalana, Catalònia and Proa and collections like the Quaderns Literaris. Leading Catalan writers, aware of the importance of this activity for redressing the cultural and linguistic balance, became translators themselves. For translations from English, two of the most important figures were Josep Carner, who translated Mark Twain, Lewis Carroll and three of the great Dickens novels, and later on, Marià Manent, who translated a large number of English-speaking poets. The first indirect translations from Russian date from 1886 thanks to the interest of novelist Narcís Oller, who in 1897 translated three narratives by Tolstoy – an author also of great interest to J. Casas-Carbó, who published two volumes of his stories in 1903. Another novelist, Joan Puig i Ferreter, re-translated Tolstoy and introduced Gorky in 1909. In the 1920s the first direct translations appeared: some of the works by Dostoyevsky, Tolstoy, Turgenev and Chekov were translated by Francesc Payarols; the politician and essayist Andreu Nin had a firm grasp of the Russian language (having spent ten years living in the country) and translated novels as iconic as Dostoevsky's *Crime and Punishment* (1929) and Tolstoy's *Anna Karenina* (1931), as well as more contemporary authors. In the German language, Goethe had a considerable influence: Joan Maragall translated four of his works between 1898 and 1910; Josep Lleonart translated *Torquato Tasso*, *Hermann und Dorothea* and *Faust*, while Joan Alavedra translated *Werther* (Llovet 1982). From Italian, there was particular interest in classical literature and specifically in Dante's *Commedia*; in 1908 a prose summary by Antoni Bulbena was added to Andreu Febrer's 15th century translation, while a partial translation by Narcís Verdaguer i Callis in 1921 and prose and verse versions by Llorenç Balanzó were published in 1923–1924 (a translation in tercets by Josep M. de Sagarra would appear later between 1947 and 1951, followed by a new version by Joan F. Mira in 2000).

Bible translation has been an important resource in the formation of many national languages. Towards the end of the 19th century and the beginning of the 20th, Antoni Bulbena, Francesc de P. Castells, Pere Marcé, Tomàs Sucona and Frederic Clascar embarked on a Catalan translation of the Bible. Spanish hegemony meant that no complete Catalan versions of the Bible were permitted until the 20th century. Once the ban was lifted, three appeared simultaneously in the pre-war period, but with very different characteristics: that of the Foment de Pietat Catalana, the Monestir de Montserrat version and that of the Fundació Bíblica Catalana (the first two translations listed were never finished) (Ferrer 2009; Parcerisas 2009). Bypassing the rigour of censorship under Franco, Montserrat published a complete work in a single volume (1970) and, more recently, has added a new, interconfessional Bible (1993) as well as a Protestant version (2000).

The task of translating has often been accompanied by a theoretical reflection on the activity in itself. The list of those who have made substantial contributions is a

long one, as seen in the anthology *Cent anys de traducció al català (1891–1990)* (Bacardí/Fontcuberta/Parcerisas 1998). Here, we give just one example. In 1938, towards the end of the war, Cèsar-August Jordana systematised his broad experience as a translator in the article *L'art de traduir* ('the art of translation'), in which he called for the compilation of an inventory of outstanding translations and, later, a history of modern translation. In practical terms, he also proposed a broad classification of translators, dividing them into those who tend to translate more literally and those more concerned with equivalence. He personally supported the latter method and insisted on the need to master the target language, which he illustrated with numerous comments and examples which are still relevant today.

6 The Franco dictatorship

It is calculated that over a thousand translations were published as books from the *Renaixença* in the mid-19th century to 1939, when the Franco regime repressed all public manifestations of Catalan culture. This censorship was ruthless and the prohibition of the publication of translations was particularly rigorous in the 1940s and 1950s. The first authorised Catalan translations during this period were Riba's *Odissea* and Sagarra's *Divina Comèdia*, which received permission in 1948 on two conditions, which were extended to other works: firstly that, given the prestige of the translators and the original texts, they could be considered 'literary creations', and not translations; and secondly, that they should be published in luxury editions with small print runs and sold at sky-high prices. After that, whether a translation was authorised or not was always highly arbitrary and depended on the circumstances, personal pressure, the censor, the type of work, and the translator, among other factors (Gallofré i Virgili 1991).

In 1962, a change in government tactics lifted the ban on translations and, from that time on, they were subject to the same controls as other books. The appearance in the same year of the publisher Edicions 62 signalled a revolution in the publication of translated books, mainly thanks to the *El Balanci* collection, dedicated to contemporary literature, *La Cua de Palla*, which published detective novels, and *Llibres a l'Abast*, which specialised in essays. In addition, in 1964 the *Biblioteca a Tot Vent* collection of the Proa publishing house set out on a new mission spearheaded by Joan Oliver. According to Francesc Vallverdú (1987), of the total works produced in Catalan in 1965, 55 % were translations. However, a few years later the situation reversed as a result of the publishing crisis and saturation of the market; in 1973, for example, only 8 % of books published were translations. The translations that were produced, however, bore the names of many of the leading writers of the time: Joan Oliver, Maurici Serrahima, Xavier Benguerel, Rafael Tasis, Lluís Ferran de Pol, Joan Sales, Manuel de Pedrolo, Maria Aurèlia Capmany, Gabriel Ferrater, Joan Fuster, Josep Vallverdú, Jordi Sarsanedas, Ramon Folch i Camarasa, among others. Some indivi-

duals began to make a name for themselves as translators: Bonaventura Vallespinosa, Josep M. Güell, Carme Serrallonga, Jordi Arbonès, to name but a few. It would seem that there was a desire to make up for lost time and re-enter the fight for the prestige and authority of the Catalan language (Bacardí 2012).

7 Recent decades

Following the deceleration which occurred in the 1970s, the 1980s saw the publication of translations pick up again and experience a new ‘phase of expansion’ (Broch 1991), representing 16.5 % of all Catalan publications in 1990, for example (Vallverdú 1992, 92). In 1981, two new collections appeared for the publication of foreign works: Les Millors Obres de la Literatura Universal and Textos Filosòfics. Joaquim Mallafrè published his translation of an iconic title, James Joyce’s *Ulysses* – an undertaking which forms much of the basis of *Llengua de tribu i llengua de polis: bases d’una traducció literària* (1991), (‘Language of the tribe and language of the polis – basis for literary translation’), the first book in Catalan devoted to translation as an academic discipline. Previously Jaume Tur in his *Maragall i Goethe. Les traduccions del Faust* (1974) (‘Maragall and Goethe. The translations of Faust’) did not limit himself to these writers, but dealt with aspects of translation as a method.

Throughout the 1980s other collections appeared which were sustained by translations, such as the Clàssics del Pensament Modern, L’Arcà, Venècies, Clàssics Moderns and a new 20th century collection of Les Millors Obres de la Literatura Universal. Segle XX (Mallafrè 2000). However, a number of these collections were suspended during the following decade.

With the turn of the century, the market was flooded with translations of global best-sellers and works by more or less *literary* authors who had achieved international success, from Dan Brown and J. K. Rowling to Paul Auster and Haruki Murakami. Similarly, a new phenomenon emerged: the translation from Spanish into Catalan, given commercial possibilities for a demand that was not “real” as far as readers understood both languages, but the plot or the writers were felt as part of a Catalan setting, and this blurred the line between *original* and *translation* (Arenas/Škrabec 2006). This mainly homogeneous market stood in contrast to that of the smaller publishers (Adesiara, Edicions de 1984, Edicions del Periscopi and Raig Verd), which opted for the recovery of unquestionable classics or the inclusion of consolidated contemporary works. In the words of Josep Marco (2000, 44), it could be concluded that, today, ‘translations continue to be highly abundant, [...] but they are further removed from nuclei of influence’ than in previous times.

On the other hand, thanks to favourable conditions, these years saw the professionalisation of translation. Now it is not only writers, either out of a sense of duty to their country or because they are the only ones the publishers trust, who take on the lion’s share of translation work. Today it is the professional translator who dominates

and is able to more or less make a living from their work as long as the publishing industry remains stable, albeit with ups and downs. This is evidenced by the establishment of specific translation sections in writers' associations, such as the Association of Writers in Catalan and the Catalan PEN. It may also be as a result of the grants awarded annually by the Institució de les Lletres Catalanes and the Institut Ramon Llull to translators and publishing houses for the translation of literary works into and out of Catalan, as well as other translation prizes that have appeared in recent years. Evidently, some of the main contributors to this professionalisation have been the six Catalonia universities which offer degrees in translation. Additionally, the Universitat Autònoma de Barcelona published the first specialised periodical mainly in Catalan: *Quaderns de Traducció i Interpretació* (1982–1992), which reappeared in 1998 under the name *Quaderns. Revista de Traducció*. Since 1995, academic research has been disseminated through the Biblioteca de Traducció i Interpretació collection, published by Eumo, which contains around twenty volumes of essays, manuals and anthologies which have contributed to the creation of Catalan translation studies (Torrents 2010).

8 Translation from Catalan

Despite the historical difficulties, there is a large body of literature produced in Catalan and numerous authors have crossed borders and been translated into other languages. Some of these date back to the Middle Ages and even to the centuries of the *decadència* in Catalan literary production. Ramon Llull is certainly the oldest and best-known author: the *Llibre del Gentil e dels tres savis* was translated into Spanish in the 14th century; *Fèlix* was translated into Spanish, French and Italian in the 15th century; the *Llibre de l'ordre de cavalleria* was translated into English from the French version by William Caxton in 1484 as the *Book of the Order of Chivalry*, and during the first few decades of the 20th century Edgar Allison Peers revived interest in Llull with the translation of five of his works. There have also been translations of the great *Cròniques*, the prose of Francesc Eiximenis, the poetry of Ausiàs Marc and, in particular, the novel *Tirant lo Blanc* by Joanot Martorell, which has been translated into fifteen languages, either directly from Catalan or from intermediary languages, primarily Spanish, which has often acted as a bridge language or even an original source.

In the 19th century the fame of some writers gave rise to new translations. Perhaps the most influential was the poet Jacint Verdaguer, whose works were reproduced in about a hundred different versions in Spanish, Occitan, French, Portuguese, Italian, English, German, Czech, Russian, Latin, Esperanto, etc. About a third of these are versions of *L'Atlàntida*. Àngel Guimerà follows closely behind with more than fifty translations of his works; his play *Terra baixa*, praised by Piscator, is the work on which the operas *Tiefland* by Eugen d'Albert (first performed in Prague in 1903) and *La catalana* by Fernand Leborne were based, and it has also served as the inspiration for several films produced in Europe and the United States; José Echegaray translated

many of his works into Spanish, with *Mar i cel* and *Terra baixa* being the most frequently published and performed in many languages. In 1886 Narcís Oller broke into the French market with his first novel *La papallona*, with a prologue by none other than Émile Zola, leading to the popularity of Oller's other work to snowball throughout Europe.

Of the 20th century Catalan writers, Mercè Rodoreda is undoubtedly one of the most frequently translated; *La plaça del Diamant* can be read in around thirty languages, with three versions in English alone and two in Spanish, Italian, Dutch and Portuguese. Other writers from the same generation who have been translated include Josep Pla, Llorenç Villalonga, Joan Sales, Manuel de Pedrolo and Joan Perucho. As far as the poets are concerned, versions of works by Josep Carner, Carles Riba, J. V. Foix, Salvador Espriu, Joan Brossa, Gabriel Ferrater and Miquel Martí i Pol are among the most notable.

Younger writers have also been exported profusely. These include works of narrative by Baltasar Porcel, Jesús Moncada (*Camí de sirga*, in twenty languages), Isabel-Clara Simó, Montserrat Roig and Carme Riera, the poetry of Joan Margarit, Narcís Comadira, Francesc Parcerisas and Maria-Mercè Marçal and plays by Josep M. Benet i Jornet. During the 1990s, the stories and novels of Quim Monzó were translated into twenty languages. Catalan literature gained a certain presence in European markets thanks, to a large extent, to certain foreign translators who acted as ambassadors, and to the internationalisation of certain publishing houses and literary agents. Sergi Pàmies and Lluís-Anton Baulenas, for example, were also translated. A surge in writing for the theatre during that time meant that work by playwrights such as Lluïsa Cunillé, Carles Batlle and Sergi Belbel were – and continue to be – performed in many languages.

The translation of Catalan literature has taken a new direction since Catalonia hosted the Frankfurt Book Fair in 2007. The event saw the presentation of 53 new German translations of authors of very different backgrounds, periods and styles: Lluïsa, Martorell, Víctor Català, Sagarra, Pla, Espriu, Pere Gimferrer, Maria Antònia Oliver, Belbel and Ada Castells. Maria Barbal, with *Pedra de tartera* (50,000 copies sold), and Jaume Cabré, with *Les veus del Pamano* (40,000), attracted critical attention and almost unheard-of public success (increased by the reception of his next novel, *Jo confesso*, translated into more than twenty languages), followed by *Pandora al Congo* by Albert Sánchez Piñol (25,000), who had already seen his *La pell freda* translated into over thirty languages (Jané Lligé 2012, 166). Two months after the fair, it was calculated that in total 'around 250,000 books by Catalan authors had been sold in Germany' (Torner 2012, 118).

Since then, Catalan literature has become much more widely known, read and translated throughout the world and, even though it has not been possible to sustain such high figures for translations – or sales – when all is said and done 'the Catalan presence can be compared perfectly well with other languages of culture' (Mallafre 2009, 16).

9 Bibliography

- Aramon i Serra, Ramon (1985), *Frederic Mistral i la Renaixença catalana*, Barcelona, Rafael Dalmau.
- Arenas, Carme/Škrabec, Simona (2006), *La literatura catalana i la traducció en un món globalitzat*, Barcelona, Institució de les Lletres Catalanes/Institut Ramon Llull.
- Bacardí, Montserrat (2002), *Notes on the History of Translation into Catalan*, *Catalan Writing* 17–18, 13–99.
- Bacardí, Montserrat (2012), *La traducció catalana sota el franquisme*, Lleida, Punctum.
- Bacardí, Montserrat/Estany, Imma (2006), *El Quixot en català*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Bacardí, Montserrat/Fontcuberta, Joan/Parcerisas, Francesc (edd.) (1998), *Cent anys de traducció al català (1891–1990). Antologia*, Vic, Eumo.
- Bacardí, Montserrat/Godayol, Pilar (edd.) (2010), *Una impossibilitat possible. Trenta anys de traducció als Països Catalans (1975–2005)*, Vilanova i la Geltrú, El Cep i la Nansa.
- Bacardí, Montserrat/Godayol, Pilar (edd.) (2011), *Diccionari de la traducció catalana*, Vic, Eumo.
- Bacardí, Montserrat/Godayol, Pilar (2013), *Les traductores i la tradició. 20 pròlegs del segle XX*, Lleida, Punctum.
- Badia, Lola (1991), *Traduccions al català dels segles XIV–XV i innovació cultural i literària*, *Estudi General* 11, 31–50.
- Broch, Àlex (1991), *Una etapa oberta: sota el signe de la traducció*, *Revista de Catalunya* 12, 156–171.
- Buffery, Helena (2007), *Shakespeare in Catalan. Translating Imperialism*, Cardiff, University of Wales Press.
- Ferrer, Joan (2009), *Aproximació panoràmica a les traduccions bíbliques catalanes de començament del segle XX*, in: Marcel Ortín/Dídac Pujol (edd.), *Llengua literària i traducció (1890–1939)*, Lleida, Punctum, 33–50.
- Franquesa, Montserrat (2013), *La Fundació Bernat Metge, una obra de país (1923–1938)*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Gallofré i Virgili, Maria Josepa (1991), *L'edició catalana i la censura franquista (1939–1951)*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Godayol, Pilar (2008), *Triplement subalternes*, *Quaderns. Revista de Traducció* 15, 41–50.
- Jané Lligé, Jordi (2012), *L'impacte de la fira de Frankfurt a Alemanya*, in: Teresa Iribarren/Simona Škrabec (edd.), *Constel·lacions variables. Literatura en la societat de la informació*, Barcelona, UOC, 145–186.
- Janés i Nadal, Alfonsina (1983), *L'obra de Richard Wagner a Barcelona*, Barcelona, Fundació Salvador Vives Casajuana/Ajuntament de Barcelona.
- Llanas, Manuel (2002–2007), *L'edició a Catalunya*, 6 vol., Barcelona, Gremi d'Editors de Catalunya.
- Llovet, Jordi (1982), *Goethe a Catalunya: 1832–1982*, in: *Goethe. 1832–1882. Antologia que la Generalitat dedica a les escoles de Catalunya*, Barcelona, Proa, 135–147.
- Malé, Jordi (2006), *Carles Riba i la traducció*, Lleida, Punctum.
- Mallafrè, Joaquim (1991), *Llengua de tribu i llengua de polis. Bases d'una traducció literària*, Barcelona, Quaderns Crema.
- Mallafrè, Joaquim (2000), *Models de llengua i traducció catalana*, *Quaderns. Revista de Traducció* 5, 9–27.
- Mallafrè, Joaquim (2009), *Traduccions del català avui. Dades i reflexions*, in: Kálmán Faluba/Ildikó Szijj (edd.), *Actes del Catorzè Col·loqui Internacional de Llengua i Literatura Catalanes*, vol. 1, Barcelona, Publicacions de l'Abadia de Montserrat, 15–36.
- Marco, Josep (2000), *Funció de les traduccions i models estilístics: el cas de la traducció al català al segle XX*, *Quaderns. Revista de Traducció* 5, 29–44.
- Murgades, Josep (1994), *Apunt sobre noucentisme i traducció*, *Els Marges* 50, 92–96.

- Parcerisas, Francesc (2009), *Traducció, edició, ideologia. Aspectes sociològics de les traduccions de la Bíblia i de l'Odissea al català*, Vic, Eumo.
- Paredes Baulida, M. (1999), *Traductors i traduccions a la Menorca il·lustrada*, in: Francisco Lafarga (ed.), *La traducción en España (1750–1830). Lengua, literatura, cultura*, Lleida, Universitat de Lleida, 79–89.
- Pla i Arxé, Ramon (1975), *L'Avenç (1881–1915): la modernització de la Renaixença*, Els Marges 4, 23–38.
- Pons, Arnau/Simona Škrabec (edd.) (2007–2008), *Carrers de frontera. Passatges de la cultura alemanya a la cultura catalana*, 2 vol., Barcelona, Institut Ramon Llull.
- Pujol, Dídac (2007), *Traduir Shakespeare. Les reflexions dels traductors catalans*, Lleida, Punctum.
- Pujol, Josep (2002), *Expondre, traslladar i reescriure clàssics llatins en la literatura catalana del segle XV*, Quaderns. Revista de Traducció 7, 9–32.
- Pujol, Josep, et al. (2004), *El àmbit de la cultura catalana*, in: Francisco Lafarga/Luis Pegenaut (edd.), *Historia de la traducción en España*, Salamanca, Ambos Mundos, 623–719.
- Sellent i Arús, Joan (1998), *La traducció literària en català al segle XX: alguns títols representatius*, Quaderns. Revista de Traducció 2, 23–32.
- Torner, Carles (2012), *Petita festa de la traducció de poesia catalana 2003–2010*, in: Teresa Iribarren/Simona Škrabec (edd.), *Constel·lacions variables. Literatura en la societat de la informació*, Barcelona, UOC, 115–144.
- Torrents, Ricard (2010), *Vers una traductologia catalana*, in: Kálmán Faluba/Ildikó Szijj (edd.), *Actes del Catorzè Col·loqui Internacional de Llengua i Literatura Catalanes*, Barcelona, Publicacions de l'Abadia de Montserrat, 361–379.
- Tur, Jaume (1974), *Maragall i Goethe. Les traduccions del Faust*, Barcelona, Universitat de Barcelona.
- Vallverdú, Francesc (1987), *Cinquanta anys de l'edició en català (1936–1986)*, in: *Edicions 62. Vint-i-cinc anys (1962–1987)*, Barcelona, Edicions 62, 111–118.
- Vallverdú, Francesc (1992), *El creixement del llibre en català des de 1975*, in: Joan Fuster, *L'aventura del llibre català*, Barcelona, Empúries, 83–98.
- Vila, Pep (1996), *Les traduccions d'obres franceses i italianes en el teatre català al Rosselló*, Revista de Catalunya 109, 103–129.

Catalan Today

Joan A. Argenter

17 Languages in Contact: A Sociocultural Approach

Abstract: Over the course of history, Catalan has come into contact with a number of Semitic and Indo-European languages. This chapter addresses the sociocultural dimension of language contact in the Catalan-speaking area. I adopt an ecological approach based on qualitative, ethnography-oriented, rather than quantitative methods. The effects of language contact can be seen in language patterns (e.g. interference), psychosocial patterns (e.g. language attitudes), and, sociocultural patterns (the community's "language economy"). Salient topics in this latter field are the social use of language (individual language choices as well as social patterns of functional distribution); interactional communicative practices (e.g. codeswitching); long-term sociocultural changes (e.g. language shift); technologies of language in the media. Nowadays the range of languages in contact has dramatically increased due to global demographic trends and immediate non-face-to-face, technology-based communication.

Keywords: language ecology, codes, code-choice, codeswitching, communicative practices, language shift, language attrition and obsolescence

1 Aspects of language contact

When approaching language contact the researcher must discern the processes leading to contact, the conditions under which language contact evolves, and the outcomes of language contact. The former processes are responsible for languages coming into contact with one another and may be demographic (e.g. migration), economic (e.g. trading), military (e.g. conquest), religious (e.g. conversion, liturgy), political (e.g. dominance, language policy), cultural (e.g. oral vs. written, technology), sociocultural (e.g. rites of passage), broadcasting (radio, tv, internet), regional and global social changes (large scale displacement, local, migrant and global languages). Long-term historical processes such as industrialisation, urbanisation and globalisation are also to be considered here. The aforementioned conditions are responsible for the evolution of language contact and its outcomes either linguistic (e.g. language mixing) or sociolinguistic (e.g. language shift). The direction and effectiveness of this evolution is proportionate to the degree of political or cultural pressure, to what extent a language is accessible or not to non-native speakers, to sociolinguistic prestige, and so on. As the mentioned instances suggest, language contact implies more often than not an asymmetric relationship, that between a dominant and a

<https://doi.org/10.1515/9783110450408-025>

receding language. Terms such as majority and minority languages are relative: reverse relations may hold between the same languages in different regions; the initial relationship may in some cases be reversed, or national minorities may happen to be regional majorities. Moreover, these concepts are somewhat misleading, so the terms dominant/receding will be used preferentially.

Such processes are external to language contact itself, whereas the conditions shaping linguistic and sociolinguistic outcomes tend to have an ideological contact-derived basis. These conditions include: public and institutional use vs. private restricted use; written vs. oral; language(s) of/in schooling and the language ideology transmitted; the ideological discourse of power; sociolinguistic prestige; language ideology conveyed by global vs. local status, and so on – also those factors that contend and restrain political pressure, such as close-knit social networks, language loyalty, cultural resistance, counter-hegemonic discourse, revival efforts, etc.

It is difficult to identify one of these factors as causing a specific outcome, with most outcomes being the result of multiple causation. Heavily relying on receding language speakers' greater or lesser success in learning the dominant language, Thomason/Kaufmann (1988, 46–57) established a typology of language-contact outcomes: linguistic, ranging from lexical borrowing to relexicalization or hybridization, as well as sociolinguistic, from language restriction to language shift or language death. For their part, ethnographers emphasise the role of intervening variables between macro socioeconomic, sociopolitical, sociocultural events and structures, and language shift or language death (Gal 1979, 3).

However, as Weinreich (1970, 1) once stated, since languages are “in contact if they are used alternately by the same persons”, the locus of contact is “the language-using individuals” – which is tantamount to saying that language contact takes place first and foremost in the individual's mind. Weinreich thinks of contact as causing bilingualism – on the understanding that this is not a dichotomous competence and behavioural choice, but one over a continuum of fluency. Bilingualism is thus thought to be a psychological fact.

Indeed, research on bilingualism and multilingualism has been undertaken from cognitive, psychological and neurological approaches, focusing on cognitive abilities, learning rates, performance tasks, attrition rates, and on the activity of brain structures and connections in the speech of bilingual individuals.

Nonetheless, every language has a public dimension, a fact which links language with social and cultural context, social structure, political status and eventually human social communication – either within the family, close social networks, workplace, religious practices or other social practices and scenes. As stated above, this is our field of concern here.

2 Catalan in contact: a sketch of a chronology

2.1 Indo-European and Semitic Languages

Catalan is a Romance language that evolved from Latin by the 8th century and is also the result of an autonomous development from an original Catalan/Occitan bundle (≈10 Early Medieval Catalan), its region of origin extending along both sides of the Eastern Pyrenees to the Mediterranean Sea. History would have it that Catalan spread southwards into the Iberian Peninsula and eastwards across the Mediterranean, while transforming into a language of culture, the vehicle of a flourishing literature, of philosophical and scientific thought, of law, notarised documents, and Chancery writings. Catalan acquired an elevated institutional status and was used in the Catalan-Aragonese Crown and beyond (≈11 The Growth and Expansion of Catalan (1213–1516)).

As well as its contact with Gallo-Romance (Occitan, French) and Ibero-Romance languages (Aragonese, Castilian), with Sardinian and Italo-Romance languages and varieties (Genoese, Sicilian, Neapolitan, Italian), Catalan came into contact with Semitic languages (Arabic, Hebrew) – one could add a brief interaction with Greek as the result of Catalan incursions in the Eastern Mediterranean. The Mudejar population in the Catalan-speaking regions used to live in communities where they retained the right to use their language – Andalusi Arabic –, religious practices and customs in Valencia; while they shifted to Catalan in Majorca where they lived under harder conditions. Jewish communities in Catalonia, who enjoyed relatively high status in the 13th and 14th centuries, used Catalan as their colloquial language and Hebrew as a literary and liturgical language (Argenter/Ferrer forthcoming). Eventually, Hebrew became – as it was the case everywhere – a *Textsprache*.

Contact with Sardinian was a consequence of the incorporation of Sardinia by the Catalan-Aragonese Crown (1323), which lasted until 1713. Eventually, after centuries of passing from one foreign power to another, Sardinia was integrated into the duchy of Savoy (1720) under the name of Kingdom of Sardinia and later into the Kingdom of Italy (1861). The town of l'Alguer (It. Alghero),¹ repopulated in 1354 by Catalan settlers, is a lively reminder of the Catalan-Aragonese domination of the island, since Catalan has survived as the common everyday local language in the town. Migratory waves of Sardinians into l'Alguer increased the number of speakers of Catalan, since immigrants needed to learn the language to earn a living – at the same time introducing Sardinian in the town. L'Alguer remains a Catalan linguistic enclave in Sardinia (Argenter 2008). Currently, however, despite legislation protecting minorities and intangible heritage, the local use of Catalan is highly endangered due to the pressure from Italian.

¹ The Catalan name of the town is in fact *l'Alguer*, the definite article always preceding the placename.

As a consequence of the treaty of the Pyrenees (1659) between the kingdoms of France and Spain, Catalonia lost its northernmost counties to France. In fact, the relationship between the inhabitants of the north and the south by means of their common language – despite its legal prohibition in France (1700) – continued to develop without interruption up until the 20th century. Universal education in France was what really led to an abrupt break in the use of language, with the First and Second World Wars also representing key turning points in the “Frenchification” of the Northern Catalans.

2.2 The core case: Catalan/Spanish language contact

It is undeniable that the significance of the language contact between Catalan and Spanish deserves especial attention. Let us briefly review the relevant historical facts.

The Catalan-Aragonese Crown was formed as a confederation in 1137 by the marriage of Ramon Berenguer IV, Count of Barcelona, and Peronella, Queen of Aragon. Catalan became the official language of the crown.

The death of King Martin the Humane without an heir in 1410 represents the end of the House of Barcelona and led to a new Castilian dynasty, the Trastámara, taking over control of the Confederation (1412).

The dynastic union between the Crown of Aragon and the Crown of Castile in 1474 (following the marriage of Ferdinand II and Isabella I in 1469) led to the progressive political and cultural subordination of the former to the latter. The Hispanic Crown – a primarily Castilian monarchy – was born, with a strong power structure and a great empire soon to come. The links between the kingdoms of the Catalan-Aragonese confederation weakened, leading to increasing awareness among the people of their belonging to a larger political space, the Spanish crown – a process that progressed at very different rates in the kingdoms of Aragon, Valencia, Majorca and the Principality of Catalonia. Since the 16th century, the political subordination and the cultural weakness of Catalan – leaving behind a period of great literary creativity in all genres – continued uninterrupted, and the situation worsened well into the 18th century under the rule of the new Bourbon dynasty in Spain. During this period, the representative and governmental institutions of Catalonia and Valencia were annihilated, both were administratively and juridically incorporated into the Kingdom of Castile, Spanish was imposed as the only official language and the use of Catalan as a public language was repressed (↗12 The Origins of Modern Catalan: Cultural and Linguistic Evolution).

However, the transmission and use of Catalan as an oral language was never entirely abandoned, though the social functions of the language grew more and more limited from the 16th century onwards. The aim of the *Renaixença* (↗13 *Renaixença*) was the recovery of Catalan as an autonomous written language – function more than form was the goal. It was the *Noucentisme*, a cultural and political movement in the

early decades of the 20th century in Catalonia that was responsible for setting a universally accepted normative written form, a modern literary language. A happy conjunction of intellectuals and politicians working together designed a programme to modernise the country, with the Mancomunitat de Catalunya as its political and administrative instrument from 1914 onwards. Its first president was Enric Prat de la Riba. Seven years earlier Prat de la Riba, then president of the Barcelona Provincial Council, had founded the Institut d'Estudis Catalans (1907), an institution entrusted with the study of Catalan in all its dimensions and with the standardisation of modern Catalan, a goal successfully reached under the leadership of Pompeu Fabra (1845–1912). The Language Reform, the Institut d'Estudis Catalans and the Work of Pompeu Fabra).

The French historian Pierre Vilar (1962) in his landmark work on modern Catalonia remarked that the conflict between a language that is taught at school but is unknown in the home and a language that is used by family but is not taught often leads people to a stress situation that compels them to make a choice. Vilar found it noteworthy that Catalonia, confronted with such a dilemma at the beginning of the 20th century, “n’a pas choisi la langue d’État, au sens où celle-ci a été choisie en France, c’est-à-dire en apparaissant comme la langue nationale reconnue de tous.” Moreover, Vilar – an Occitan – considered that the failure of the *Felibrige* was due to its inability to engage large swathes of society in the movement, contrary to the success obtained by the Catalanist movement: “L’emploi du catalan, comme signe d’opposition, de refus envers Madrid, n’avait de chance de triompher, dans une société divisée en classes très hiérarchisées, que si la classe dirigeante encore puissante participait à cette opposition et à ce refus” (Vilar 1962, vol.1, 162–163).

3 Effects on the social use of the languages

3.1 Individual language choice: bilingualism and multilingualism in a societal multilingual setting

A common outcome of language contact is bilingualism or multilingualism, depending on social conditions: it may be that one of the groups in contact is bilingual itself or is the bearer of a sacred language. Nowadays increasing population mobility, omnipresent mass media, the spread of world languages and the emergence of English as a global language are all emerging trends. As a consequence, the number of multilingual individuals is growing, either organically or the result of targeted policies such as those of the EU.

Speakers confronted with such an environment are continually forced to make choices. By means of language choice speakers position themselves with relation to others socially and culturally. They may maximise or minimise social and cultural cleavages; accept or challenge their higher or lower position in society; maintain

language loyalty; express power or solidarity. They may choose to present themselves as members of an ethnolinguistic group – e.g. as Catalan-speaking Sardinians – or as citizens of a nation-state – e.g. as Italians –, mark their identities, negotiate the language of interaction or take a political stance.

Language choice is not random. It follows certain locally defined patterns. Fishman's (1965) approach regarding "who speaks what language to whom, when and where", however, failed in that it only offered a synchronic, consensual, ahistoric explanation – a crucial question was lacking: why? Or even: what for?

An alternative analysis was developed by Aracil (1982) on the basis of a "language conflict" perspective: language choice in contexts of subordination leads the speakers to make choices rooted in historical and ideological struggles whose final outcome ends up being language shift unless a process of "language normalisation" is launched to reverse this trend.² Objectively, not only top-down policies but everyday language use and interaction define a space "where the structures of oppression and linguistic structures articulate through local practices – a space where dominance and resistance emerge through language use patterns" (Hill 1993, 69).

Fishman (1972) and Aracil (1977), however, shared a sociological bias, agreeing to make "domains" of language behaviour a key concept and analytical instrument for research on language choice as conditioned by a constellation of factors intervening in communication (participants' roles, topic, place and time). The expected choice of language in each domain is norm-governed. While this approach might face criticism for being too mechanistic, synchronic and consensual, Aracil stressed the intentional component of interaction, which may challenge the prevailing norms as well as cause them to change – as in reversing language shift (Fishman 1991).³

An ethnographically-construed analysis of communication (Hymes 1972) distinguishes the "setting" and the "scene" where a "speech event" takes place as well as the role of "participants", the sequence of "speech acts", the communicative "norms", the "key", the "instrumentalities", and the "goals" of communication. These factors are defined in sociocultural terms. A room is a physical delimited space; meanwhile, a classroom is a sociocultural space where a speech event such as a lesson takes place. Teacher and student are relevant participants' roles in this setting and their mutual relation implies specific rights and obligations. One usual instrumentality is the standard language, intended goals are instruction and education, and specific genres include face-to-face questions and answers, dictations, translations, presentations, examinations. For a long time in the Catalan-speaking area in Spain, the expected

2 "Language contact" and "language choice" are facts, "language conflict" is a construct. "Language normalisation" has been a commonly used term among Catalan sociolinguists, language planners, language activists and politicians, to the point that it has now become a legal term as well (e.g. in 1983 the Catalan Parliament passed the Act of Language Normalisation).

3 Aracil's "language normalisation" and Fishman's (1991) "reversing language shift" (RLS) somehow refer to the same process or very similar processes.

mandatory language in this specific type of setting was Spanish – to the unfortunate exclusion of Catalan. The shared experience of older generations still bears witness to this policy, while the experience of those in their forties and below is varied (↗20 Teaching and Learning of Catalan).

For a long time, the choice of Catalan or Spanish in everyday encounters in the Catalan-speaking area in Spain followed a well-defined pattern. Although Catalan and Spanish had been in contact for centuries, bilingualism among the Catalan population was not widespread until the second half of the 20th century. Moreover, as a rule, immigrant Spanish-speakers living in Catalan-speaking areas had long remained monolingual (↗18 Language Demography). Bilingual Catalans' language-choice was tailored to their interlocutor's language. According to an internalised norm – 'self-hatred' (Ninyoles 1969, 74–84) or "learned helplessness", in psychologists' jargon –, Catalans used Spanish to talk to non-Catalan speakers, mainly Spanish-speakers, even if these individuals understood Catalan and the Catalans were aware of this – hence the social function of Catalan language shrank. As much as this was true, Catalan became an emotional and efficient symbol of the in-group and therefore of the identity of its members. For Catalans, to speak Catalan meant 'to be one of us'.⁴

Woolard (1989), who studied the situation in Barcelona in the early 1980s, described the norms that guided language-choice and their clear-cut ethnolinguistic significance, while suggesting at the same time that this choice reflected a class frontier too.

Since the recovery of self-government in Catalonia in the early 1980s, linguistic and education policies were implemented to promote a balanced knowledge of both languages throughout the region, i.e. to create a bilingual citizenry.⁵ In this way, the social functions of Catalan increased to an extent, since it served as a form of inter-ethnolinguistic communication – and to this very extent its earlier emblematic character as a symbol of the in-group decreased. Indeed, well before this functional recovery of the language, two phenomena appeared as a result of earlier grassroots activity among organized citizens: on the one hand, the knowledge and use of Catalan somehow spread among non-native speakers; on the other hand, Catalanness was no longer necessarily linked to or expressed in Catalan.

Earlier public campaigns launched by the Catalan government were addressed to promote so-called 'passive bilingualism', i.e. language divergence in conversations, with each speaker speaking in his or her own language. Overtly, the goal of this policy

4 This is still observed nowadays by Woolard (2016, 67): "To be sure, one can become authentically socially Catalan by mastering and using the Catalan language." Or: "Someone who speaks Catalan habitually in daily life is for most purposes taken to be Catalan."

5 This policy was a means to achieve one central goal: an "integrated society", that is to say, to avoid social segregation on the basis of language. Obviously, "integration" is multifactorial and bi- or multilateral.

was to break the internalised norm that guided Catalans' convergence with Spanish-speakers. The message being covertly conveyed was that there was no language conflict at all in Catalonia. This policy, however, cannot be evaluated as successful or realistic. Later on, the effect of schooling would help in a significant way, inasmuch as it provided access to the local language for non-locals.

Once their linguistic "learned helplessness" was partially eliminated, speakers of the Catalan ethnolinguistic group had the opportunity to construe an interaction with a member of the other ethnolinguistic group in terms of an inter-ethnolinguistic relationship or in terms of an interpersonal relationship. If in adopting a language-loyalty strategy or a power-contending stance, the speaker construed the interaction in the former terms, he or she tended to follow a divergent pattern, while if construed in the latter he or she would converge with the interlocutor in their choice of language. However, a few years later Boix, who studied the behaviour of groups of young people in Barcelona in the late 1980s, showed that the clear ethnolinguistic barrier of previous generations had become more permeable among younger generations, and that interpersonal choices superseded ethnolinguistic belonging (Boix 1993, 211): in a way language use was about to be "de-ethnicized" or "depoliticised".

3.2 Codes and the struggle to define what counts as language in the domain of the media

It was to be expected that in such new conditions negotiating the differentiation and contextualisation of "codes" – what counts as the "legitimate language" in a determined context – would become extremely relevant, underlying social cleavages. Apparent depoliticisation coexisted with political controversy over language allocation in the public arena, stressing and reinforcing language ideologies, in fact (↗23 *Language Ideologies in Society*). Indeed, controversies arose not only between ethnolinguistic groups, but internally as well.

"Language" then is not a given, but is subject to struggles over its definition. The conflict between "light" vs. "heavy" Catalan (Pericay/Toutain 1986) in the mid-1980s is a case in point. These labels were intended to refer to – or make emerge – distinct codes in conflict, since the supporters of "light" Catalan challenged the literary norm as well as the authority of the Institut d'Estudis Catalans. Spearheading this controversy were mass-media language advisors, a new professional class that sprang up with the recovery of Catalan in public life under self-government and seeking to institutionalise their position under the shelter of the media. Certainly, the new socio-political conditions demanded new language resources to fulfil the new communication needs of Catalan society in the area of print, audio and TV media (↗21 *Catalan in the Mass Media: The Rise of Stylebooks*) – all this while the major achievements of schooling in Catalan were still to come and the majority of adults had been schooled in Spanish. However, the controversy also had an ideological bias (Kailuweit 2002), both

in the field of language itself – advocating a non-puristic usage – and in the struggle for the power to define what should be the legitimate code in that domain – indirectly conditioning what any standard variety should be like. This example illustrates that “codes” do not convey mere varieties, but interested ideological positions.

4 Effects on the communicative forms of language use

4.1 Codes, codeswitching and double-voicing

Codeswitching is defined as the use of two languages or two codes in the same communicative event between speakers who are somehow bilingual. As usually understood, then, codeswitching implies the recognition of previously coexisting codes and their interplay within particular speech events in speech communities. Hence the assertion that each combined linguistic stretch keeps to its own grammar. These common premises have been qualified inasmuch as codes may emerge in an ongoing interaction out of the co-occurrence of verbal resources extracted from different sources, not necessarily from two bound language systems. Thinking in terms of language resources instead of language systems has proven to be an important shift in the search for a social foundation for dynamic linguistic facts.

The phenomenon of codeswitching, first identified as such by Jakobson (Jakobson/Halle 1956), was pointed out by Weinreich (1970, 73 and 68–69) in his classic monograph on language contact. Weinreich interpreted language switch in terms of “deviation from the norm”. Also, he held the structural principle that any utterance belongs to one language, and that this language is determined by grammar, not lexicon. In the end, both authors thought of codes as languages or at best language varieties.⁶

Gumperz (1982) was the first to conduct systematic research into codeswitching and put it firmly on the sociolinguistic agenda. He argued that codeswitching should not necessarily be understood as incomplete mastery of the language. Rather, he sought to find out what kind of social meaning was being conveyed by alternations between linguistic codes (Blom/Gumperz 1972) and what meaningful inferences were being made. He continued to think of “code” essentially as a language or a language variety. Gumperz realised that codeswitching should be approached looking at both its formal linguistic features and its communicative functions in different social settings. However, the search for particular or universal principles of the “grammar”

⁶ Jakobson, always concerned with the idea of a dynamic synchrony, was referring to differences in the speech of older vs. younger generations.

of codeswitching is neither conclusive nor extremely insightful, and will not be considered here. Instead, functional, interactional and language as social practice approaches prevail in the following.

Ultimately, codeswitching implies that speakers involved in the give-and-take of everyday life recognise and negotiate mutual rights and obligations. Where codeswitching is more than compulsory language choice, interchanges are open to a range of meanings, styles and reciprocal stances.

As significant as the existence of codeswitching in a community's language economy is its non-existence. It is plausible that older generations of Catalans, although they shifted their language choices according to the interlocutor, did not engage in codeswitching to a significant degree, except in certain genres, such as proverbs or jokes, and language routines, such as counting or listing. Be that as it may, the loss of these expressive resources was an indication that language attrition (cf. 5.4) was lurking, with Spanish proverbs replacing Catalan ones.⁷

Two types of codeswitching were identified fairly early on: situational and metaphorical codeswitching (Blom/Gumperz 1972). The former redefines a situation in terms of the roles or the participants' mutual rights and obligations – raising shared expectations among the participants. Instead, metaphorical codeswitching hinges on intentional styling of speech “where alternation enriches a situation, allowing for allusion to more than one social relationship within the situation”.

In later work, Gumperz abandoned this distinction and emphasised the descriptive concept of “contextualisation cues”: those linguistic and paralinguistic features aimed at setting the context on whose basis an utterance or speech is to be construed in the ongoing interaction; codeswitching side by side with intonation, modulated voice, whispering, voice raising or lowering, lexical choices and others are devices to signal intentional meaning and alignments. Nonetheless, “if we shift our focus away from the code juxtaposition itself to the inferences that it generates” (Rampton 1998, 303) the distinction continues to be useful. The point is ease of inference: the interlocutors' ability to access the definition of new contexts of use is routinely grasped in situational codeswitching, while not straightforwardly so in “figurative codeswitching”.

4.1.1 Early identified functions of codeswitching

The following are among the main functions of codeswitching identified in the early stages of research (Gumperz 1982, 75–84):

⁷ The target of research on codeswitching has mainly been younger generations – this became relevant with the spread of Catalan and bilingualism through schooling as well as with the arrival of new migrants speaking multiple languages (↗20 Teaching and Learning of Catalan; ↗18 Language Demography).

a) Quotation:

A storyteller narrates *Little Red Riding Hood* to small children. The narration flows in Catalan until it comes to the characters' voices which trigger a switch to Spanish. This is actually not required by the well-known story, but codeswitching emerges in moments of emotional involvement – or else it reflects the fact that whoever is playing the role of storyteller learnt the story from a Spanish-speaking source.⁸

... *Llavors la Caputxeta Vermella va dir a qui creia la seva àvia: –“Pero, abuelita, que dientes más largos tienes...” –“¡Son para comerte mejor!”, va dir el llop, i d'un salt va sortir del llit i s'hi va abraonar al damunt...*

‘...Then Red Riding Hood said to whom she thought was her grandmother: – “But Grandmother, what big teeth you have...” –“The better to eat you, my dear!”, said the wolf, who leapt out of the bed and pounced on her...’

Traditional narratives are populated with different characters, fostering the concurrence of multiple voices and changes of footing (Goffman 1981), hence codeswitching, voice modulation,⁹ and other contextualisation cues.¹⁰ This behaviour is not unique to the telling of stories or myths, but occurs in everyday conversation:

– *En Jordi va guanyar el concurs de l'Ajuntament perquè tenia el suport del batlle*
– *Ja se sap, com diuen els castellans, quien a buen árbol se arrima...*

– *‘George won the City Council’s call for tenders because he was given support by the mayor*
– *As we all know, as the Castilians say, whoever latches onto the shadiest tree...’*

In the second turn, the speaker expresses obvious agreement and offers an explanation of the facts based on everyday life and experience, on “folk wisdom” as expressed

⁸ Catalan in italics, Spanish in normal script. Hereafter, italics are used for Catalan – whatever the dialect –, normal script for other languages (Spanish, French, Italian). For convenience, conventional orthography is used in transcribing oral texts, occasionally with minor modifications.

Examples come from the literature or from recordings made by the author and collaborators, and the source is indicated for each fragment – otherwise these come from interactions casually observed by the author, as is the case in this first example.

⁹ Voice modulation or accent mimicking often go hand in hand with codeswitching. Storytellers often impersonate characters and mimic their (imagined) voices, as the speaker did when playing Red Riding Hood’s voice and the wolf’s disguised voice impersonating the Grandmother. The speaker superimposed codeswitching on voice.

¹⁰ It is plausible that the codeswitching in *Red Riding Hood* is the result of its learned transmission from Spanish sources. By contrast, codeswitching is unusual in a story such as *La rateta que escombrava l’escaleta* (known as *The vain little mouse* in English), since it has a stronger tradition of oral transmission, as if belonging to a traditional Catalan repertoire of stories.

by means of traditional “texts”,¹¹ in this case a Spanish proverb. The meaning of this text is not to be construed in terms of propositional semantics, but based on pragmatic and cultural knowledge and inference. Moreover, uttering a proverb works as quotation. As is the case here, often the speaker “reproduces” only the first line of the proverb, since the addressee is supposed to complete the “text” from shared background. What the speaker means is not explicitly asserted, but inferred from the proverb and its context-embedded history. This is an instance of metaphorical codeswitching and illustrates indirect speech.¹² The parenthetical remark *com diuen els castellans* – apparently the source of the quote – is a common disclaimer indicating that what follows is a Spanish word or phrase or text – not necessarily a proverb.

Both codeswitching instances, the one in the story and the proverb, minimise the “intertextual gap” between the “original” source and the “copy” (Bauman/Briggs 1990) as the latter replicates the former – both the dialogue in the story and the proverb itself are highly fixed – and go unchallenged by the speaker.¹³

Quotation codeswitching is found in Catalan/French alternation among Catalan gypsies in France (cf. 4.1.3). The ongoing conversation addresses marriage and wedding traditions among gypsies. The speaker reports her experience as an “unclean” bride:¹⁴

–*Vam fer la festa e après ell és anat a la seua casa i jo som quedat aquí. I els paios me deien: “Où est ton mari?” E jo tenia vergunya, deïi: “Ils sont partis au café avec ses frères”. La noia se casa sense sapiguer res, sense sapiguer res!* (Casanova 2016, 331).¹⁵

–*We had the feast and then he went home and I remained here. And the paios [non-gypsies] asked: Where is your husband? And I was ashamed, I said: “They’ve gone to the Café with their brothers.” A girl marries and she does not know anything, she does not know anything!*

Obviously not all quotations are subject to switch, as is shown in the next case – contributed by an old woman interviewed in l’Alguer. After a direct answer to the question asked by the interviewer, the interviewee constructs a narrative containing

11 In the sense of Silverstein/Urban (1996, 5): “meaningful building blocks of shared culture” or verbal and ideological resources that are mobilised through social practices.

12 The complete proverb is *Quien a buen árbol se arrima, buena sombra le cobija* (‘He who snuggles up to good tree, good shade shelters him’) and is equivalent to the English expression ‘It helps to have friends in high places’.

13 Codeswitching is always locally-bound. Hence, not all stretches marked as Catalan belong to the same “code”. Side by side with standard Catalan and peninsular dialects, instances from Catalan varieties in contact with French, Sardinian and Italian are presented. Historical loanwords from these languages – and from Italic dialects – in such cases are to be considered code elements when these are terms commonly used by the speakers in their otherwise Catalan code.

14 She got her period on the wedding day.

15 The term *paios* designs non-gypsies. The word has been incorporated as a traditional borrowing into common Catalan.

two embedded reported speech events: Gavino, the speaker's late husband, recounts to her (first speech event) a casual interchange between youngsters (second speech event) that he has just witnessed. The speaker, who is speaking in local Catalan with the interviewer, switches to Italian when quoting her husband and back to Catalan. When the speaker's narrative reaches the point in which the interchange between the youngsters is reproduced no codeswitching arises. After so many years, their voices are still heard in local Catalan.¹⁶ The fragment illustrates the speaker's metalinguistic and metapragmatic awareness (Silverstein 1992).¹⁷

– *I dels minyons d'avui cosa ne pensa?*

– *Mah dels minyons d'avui ja hi ha minyons que ja són educats, ma ja n'hi ha propiu que no va bé a diure. La dona és pitjor de l'home. Peggio che mai. Gavino un dia n'és vengut assustat. M'ha dit: "Niné, ma é una cosa! Io mi sono fermato a guardare questa ragazza", i xint d'escola, aquell que l'aquidrava, i ella li dieva: "eh, eh...". A l'últim s'és girada – excusa de l'expressió – i li ha dit: "cosa catzo vols?" Una ragazza ... i Gavino, si com era delicat, hi és davallat dels codus! Ma cosa? Ma jo los companyons meus que ixivem. Jo ja ixiva, ma a la passajada, mos encontraven tra amics, tra companyones, ma ningú no se permissiva de diure un frasari del gènere; ma oggi è uno schifo. No, no m'agrada a mi! (SPEAKER_046_W74).*

– *And what do you think about current day youngsters?*

– *Well... Nowadays, of course, there are youngsters who are polite, but others do not deserve the name. Women are worse than men. Worse than ever. One day Gavino came back alarmed. He told me: "Nine, you won't believe this! I stopped to look at a girl" outside the school a boy calling her and saying: "eh, eh...". At the end she turned her face and – excuse the expression – she said: "What the fuck do you want?" A girl...and Gavino, polite as he was, looked the other way! But what? Look, I, my friends and me we used to go out. I did go out, of course, for a walk, we used to gather together with friends, but none of us would dare to talk like that; but nowadays it is disgusting. No, I don't like it.*

b) Evaluation, objective statement, personal involvement, reiteration:

This same fragment illustrates other functions of codeswitching: evaluation of the message and/or personal involvement, conclusive argument, reiteration. The speaker ends her answer to the interviewer's question by providing an evaluative statement in Italian that almost reiterates – and reinforces – her final conclusion in Catalan: *La dona és pitjor de l'home. Peggio che mai*. From this point forward the speaker con-

¹⁶ Dorian (2014, 329–346) distinguishes “realistic” from “unrealistic” codeswitching, according to whether this matches the language of the original speech event or not. Apparently, here the speaker conforms to a realistic pattern – Gavino used to speak Italian with her, while we can assume that the words of the young girl were uttered in local Catalan. However, no conclusion as to a general pattern can be drawn due to scarce data.

¹⁷ Extracts whose source is (SPEAKER ...) come from published or unpublished material recorded by the author. Conventionally, each speaker is indexed with a code number and an abbreviation for sex and age. Thus (SPEAKER_046_W74) stands for a 74-year-old female speaker.

structs a narrative to confirm her point. After reporting the girl's swearword, the speaker makes an entry in her own narrative, recovers her own voice and exclaims, switching to Italian: *Una ragazza!* Her stance is a common sexist one: girls are not expected to speak that way. The last evaluative codeswitching appears when contrasting the behaviour of her generation and that of the youth nowadays, in the form of both an "objective" statement in Italian and a subjective conclusion in Catalan: *ma oggi é uno schifo. No, no m'agrada a mi!*

c) Addressee specification:

–La lengua...*la llengua materna de mosatros és castellà, de ma germana i meua, perquè mos pares, entre ellos se parlen valencià però a mosatros, des de xicotetes, mos han parlat en castellà [...]* Pero, Marc, siéntate que nos vamos. *I este també, después del castellà, està ensenyant-se en el valencià* (Montoya 1996, 187).

–The language...*our mother tongue is Castilian – my sister's and mine –, since our parents speak Valencian with each other, but they spoke Castilian to us, from childhood [...]* But, Marc, sit down, we are about to leave. *And this one too, he has been raised in Spanish, but now he is learning Valencian.*

The speaker, from Alacant, is speaking Catalan with the interviewer about the language received from her parents, that is, Spanish. At a certain moment, she turns her gaze to her child and addresses him in Spanish – except for his name: *Marc*, a Catalan form –, the language identifying her new addressee and plausibly evoking ongoing language shift.

d) Discourse markers:

The field of discourse markers is particularly sensitive to codeswitching. It is usual to hear responses such as that seen in the second turn below, initiated by the Spanish lexical item *bueno* (Cat. *bé*).

–*Pare, podem mirar dibuixos animats? Va, si us plau...*
–*Bueno, però només un capítol.*

–*'Dad, can we watch cartoons? Can't we? Please...*
–*Okay, but just one episode.'*

Within these markers tag codeswitching stands out. A Spanish-speaking tenant farmer talks with the Catalan owner of the farm:

–*El otro día íbamos a recoger las almendras, oi? Pero les eché un vistazo y todavía están verdes, así que dije dejémoslo para más adelante, oi?*

–*'The other day we went to collect almonds, didn't we? But I had a look at them and they're still unripe, so I said let us leave it for later, didn't I?'*

Oi? is a very common Catalan discourse marker,¹⁸ always at the end of an utterance: a tag question usually implying the addressee's acquiescence.

In the following a French Catalan gypsy woman is talking (cf. 4.1.3). At the very beginning she reproduces a dialogue without language alternation. Instead, two tag switching tokens appear (*ça y est* and *voilà*, closing the sentence and enforcing its truth). Other "insertional" switches (cf. below) appear too (*au contraire* – similar to the previous one, but maintaining syntactic connection –; *au mois de Mai*):¹⁹

–“*Vinc mangar [demanar] la noia*”. –“*Bueno sí, gafa la noia*” i *ça y est. Com això de la roba de casament, el mocador e la mitat de la festa és per nusaltres, l'ampolla, el menjar, l'orquestra, la sala..., tot això, i el català és au contraire. Guaita, la meua noia petita, la meua nêta, que se casa au mois de Mai, el seu sogre, paga res, perquè és el seu noi, i som nusatrus que paguem, voilà* (Casanova 2016, 340).

–“*I've come to ask for the girl*”. –“*Well, yes, take her*”...and that's all. As for the wedding clothes, the handkerchief and half the celebration are paid for by us, the bottle, the food, the orchestra, the dancing-room..., all these, and the Catalan is the other way round. Watch it! My little child, my granddaughter, who is about to marry next May, her father-in-law doesn't pay anything, because the groom is his son, and we are the ones who pay, that's all.

Next we see a blend of tag switching – though the item is not a discourse marker – and “reiteration” or “doubling”.

Sempre xerràvem u mallorquí, siempre (Montoya/Jofre 2002,786).²⁰

'We always spoke Majorcan, always.'

Tag codeswitching closes the utterance and reinforces the argument, a conclusion that expresses the speaker's stance relative to the assertion, envisaging the interlocutor's acceptance of it.

Another typology distinguishes insertional vs. alternational codeswitching (Muysken 1997; Auer 1999). In the former the speaker inserts a lexical item or a short phrase of one code into the speech flow in another code and immediately returns to the former:

–*Avui a l'escola ens han dut a visitar l'Aquàrium i hi he vist els caballitos de mar. Neden drets!*

–“*Today the school took us to visit the Aquarium and I saw the seahorses. They swim standing!*”

Instead, a return after the switch into the previous language is not predictable in alternational codeswitching. In some cases, this type yields a strategy of neutralisa-

18 *Oi* is a form related to Occ. *oc* (also the medieval Catalan form) and Fr. *oui*.

19 Note the syntactic interference of French in the initial verbal periphrasis (it lacks the preposition *a* in between *vinc* and *mangar*).

20 *u* is a form of the article in the obsolescent Catalan of the third generation descendants of Majorcan migrants in Argentina (cf. 5.4).

tion of identities – the speaker presents himself not as member of one group or another, but of both, marking his bilingual identity and double sense of belonging. This strategy tends to indicate a noticeable ideological flavour. The following is a paradigmatic and memorable example:

Gracias *Barcelona*. Gracias Cataluña. Gracias España. *Ho heu aconseguit. Aquests han estat sense cap mena de dubte els millors jocs de tota la història olímpica* (applaudiments). *En nom de tots, Comitè Organitzador, Ajuntament de Barcelona, Generalitat de Catalunya, Gobierno de España, Comité Olímpico Internacional, Consejo Superior de Deportes y numerosas empresas nacionales e internacionales han hecho posible este gran éxito... Barcelona no serà la mateixa en el futur y tampoco nuestro deporte después de las grandes victorias obtenidas [...]* (Samaranch's speech at the closure of Barcelona Olympic Games 1992; Boix 1993, 62–63).²¹

'Thank you, *Barcelona*. Thank you, Catalonia. Thank you, Spain. *You got it. No doubt that these Games have been the best in all Olympic history* (applaudes). *In the name of everyone, the Organizing Committee, the City Council of Barcelona, the Government of Catalonia, the Spanish Government, the International Olympic Committee, the Sport Superior Council and many national and international corporations that have helped to make this great success possible... Barcelona will not be the same in the future neither will be our sport after the great victories achieved [...].'*

As some of the aforementioned examples suggest, sometimes it is not easy to establish a clearcut boundary between codeswitching and borrowing – the inserted unit *caballitos de mar* may be a case in point. It is intentionality rather than automatism that characterizes codeswitching. *Caballitos de mar* may fill a lexical gap in the mind of a Catalan speaker,²² in which case it should be counted as a borrowing. However, the child may have heard the expression from another child at school and is echoing the latter's voice, making this an example of codeswitching.

Crucial to our discussion is double-voicing, that is, the speaker's intentional use of another's speech, investing it with a new meaning while not effacing the meaning that it has on its own. This is present in some of the examples given up to this point and in the following.

4.1.2 Heteroglossia, double-voicing, dialogism

A different way to approach codeswitching is to consider that speakers confront "heteroglossia" in society, make choices out of available verbal resources, and construct "voices" (Bakhtin 1981). They create their own voice and appropriate, mimic or

²¹ Barcelona-born José Antonio Samaranch was at that time the president of the International Olympic Committee. Boix explains the alternation in terms of the content connotations of each stretch (local vs. national/global).

²² Or it may reflect her linguistic insecurity: the Catalan equivalent is *cavall marí* or *cavall de mar* or *cavallet de mar*.

challenge others' voices, specifically, ways of speaking that represent particular interested positions and identities (Hill 1993, 69). These voices interact with each other dialogically, utterances responding to utterances, thus both creating and accumulating a history of context-bounded meaningfulness, the result of which is discourse "polyphony" (Bakhtin 1981).²³

Thus, in the following instance, through the participant's personal voice, the anonymous voice of public advertising reverberates in Patrícia's turn below – or else, she is mimicking this voice, which entails language alternation, before returning to Catalan to close her turn and reinforce her argument:

Patrícia: –*És una escola de pago i el curs val unaa- tot el any. Són unes cent mi- ... quasi dos-centes mil peles ... lo pagas a codo- cómodos plazos dee ochoo o diez mil pelas al mes ... però són deu bitllets* (Pujolar 2001, 192).²⁴

Patrícia: –*'It is a private school. And the course costs a-, for the whole year, it's about a hundred thousand two hundred thousand ptas. You pay it in cof- comfortable instalments oof eight or ten thousand pta. per month but, but it's ten papers!'*

Double-voicing – managing heteroglossia, stressing dialogism and contributing polyphony – is at the heart of codeswitching as a contextual "reaccentuation" mechanism (Bakhtin 1981, 339; Pujolar 2001, 30) that provides an interpretative frame. Another's voice appears in a new context, thus acquiring a new meaning or soaking up new nuances, from authority to irony.

In the following Spanish utterance, the speaker switches to Catalan, replicating the former president of the Catalan Government Jordi Pujol's common saying when he did not want to answer a question asked by a journalist or another politician or any person asking embarrassing questions:

–Los catalanes sois muy listos... Siempre estáis metiendo prisa para que se os den respuestas inmediatas, pero si os preguntan a vosotros y no os gusta la pregunta, ya se sabe, *avui no toca*.

–'You Catalans are smart people... You are always demanding immediate answers, but if you are questioned and you don't like the question, as we all know, *no comment for the moment*.'

The Spanish speaker appropriates Pujol's well-known response, at the same time presenting it as if it were a common utterance of all Catalans – not just Pujol – while actually arguing with a small group of Catalans. Instances of codeswitching are salient in terms of the co-presence of voices and the speaker's metapragmatic and metalinguistic

²³ This approach has been perceptively elaborated and empirically explored by Pujolar (2001) in his research on two urban cliques in Barcelona.

²⁴ Pujolar's extract is adapted to conventional spelling. Traits of orality represented in common spelling survive, such as non-traditional Catalan pronunciation – lack of liason (*el any* for *l'any*) –, hesitations, self-repairs, pauses (undifferentiated to length).

guistic skills in appropriating other's voices – even if appropriation by itself does not necessarily entail codeswitching.

4.1.3 Codeswitching and situated language practices

The study of language-in-context is intended to help us to understand how language works in the social world speakers live in, but also to gain insight into this world, of its order and conflicts. A step further in the analysis of codeswitching is to move away from the mere adjacency of codes towards linguistic situated practices and assume that these are actually part of broader social practices that construct social reality. Relevant questions in this area of study touch on the distribution and circulation of verbal resources and the consequences for individuals' life trajectories and for the construction of legitimacy. Before considering a case in point, I must bring up Casanova's research on Catalan gypsies in France.

Casanova (2016) has opened a window into a previously hidden or unknown social reality: the great number of Catalan gypsy communities spread across France.²⁵ Their members identify themselves as Catalan gypsies and assert that they speak *Gitano* ('Gypsy') or Catalan – that is, a peculiar variety of Catalan, indeed. At one point in his fieldwork, Casanova asks a group of Catalan gypsies in Montalban about the origin of families in the community, triggering claims of the gypsies' Catalanness (vs. "Spanishness") as well as of the language they used and eventually shifted to. Everybody gets involved in both debate and gossip about their close and not-so-close relatives, about their origin and the origin of their ancestors from Spain. When it comes to regional identity, marriages and language shift, one of them says in Catalan:

–No sem catalans, sem andalusos i aragonesos, vam vindre en França, se van casar con catalans d'aquí i vam parlar aixina (Casanova 2016, 595).

–'We are not Catalan, we are Andalusians and Aragonese, we came to France, they got married to Catalans settled here and we ended up speaking this way.'

Someone wants to be specific:

–Se van casar con los catalanes i vam parlar aixina, català (Casanova 2016, 595).

–'They married Catalans and we ended up speaking this way, Catalan.'

²⁵ Besides the communities settled in Perpignan and other villages in Northern Catalonia, Catalan gypsy communities were recorded in towns like Carcassonne, Montpellier and others in the Mediterranean arch. Casanova's finding goes much further, in quantitative human terms (number of communities and individuals), as well as territorial, social, cultural and linguistic terms.

The researcher insists on enquiring which families married into which Catalan families, and a man in the group says:

–Ellos, los Flores, *se van casar con dones catalanes, i és per això que molts parlen català: Batistes! Los Batistes parlen català. Si no, ellos no tienen res de catalans. Són andalusos i aragonesos* (Casanova 2016, 595).

–‘They, those Flores, *married Catalan women, and that’s the reason why many of them speak Catalan: Batistes! Those Batistes speak Catalan. Otherwise, they [Flores] have’nt a bit of Catalan. They are Andalusians and Aragonese.*’

They are talking about earlier generations, whose biographies crossed with those of other Catalan-speaking gypsies already settled in France. The last speaker answers to a repeated request. Finally, family names are identified. In a crucial point, he questions or denies the Catalanness of the Flores’, except as something acquired second-hand by way of the Flores’ marrying into the Batiste family. What is at stake is the legitimacy of the Flores’ claim of group belongingness. He makes the statement iconically switching into Spanish and back to Catalan in the negative quantifier *res* ‘nothing at all’.

4.1.4 Bivalency and neutralising strategies

“Bivalency” (Woolard 1999) refers to the simultaneous membership of a given linguistic segment in more than one linguistic system in a contact setting. Strategic bivalency is a language user’s deliberate manipulation of such segments – aiming, for instance, at presenting him or herself as a bilingual rather than a member of one linguistic group. Language choice follows a both-and rather than either-or pattern. This may be conceived as a “neutralisation of identities” strategy. As an example: in the 1980s antimilitary groups in Catalonia promoted the slogan *la mili no mola* – the sentence may be perceived as both Catalan and Spanish.²⁶

Nonetheless, the strategy may pursue similar but not identic aims. J. V. Foix, one of the greatest Catalan poets of the 20th century, ran a patisserie in Barcelona. He managed to avoid linguistic prohibition and censorship in both his poetry and his shop. The main sign identifying the cakeshop was CONFITERIA BOMBONERIA FOIX

²⁶ *Mili*, a truncated form, stands for ‘military service’ and *mola* comes from youth slang and means ‘excites’: “military service doesn’t excite”. Catalan/Spanish bivalency in oral language is always relative, because of either phonological or phonetic properties, e.g. the open-mid/close-mid vowel distinction of Catalan and the vowel reduction rule of Eastern Catalan. In *la mili no mola*, the sounds affected would be those represented by the letters *o* in *mola* and *a* everywhere: Cat /ɔ/ vs. Sp. /o/, and Eastern Cat. [ə] vs. Sp. [a].

de SARRIA.²⁷ Both words are bivalent (Catalan and Spanish), unlike Cat. *pastisseria* and Sp. *pastelería* – the latter being the common Spanish sign for a cakeshop. The difference between the former pair in the respective languages is phonetic (because of the vowel reduction rule in Eastern Catalan: [kamfitə'riə] vs. Sp. [kamfite'ria]) and orthographic. In Spanish, but not in Catalan, these words are written with an accent mark on final *i*, i.e. *confitería*, *bombonería*. However, it is common not to accentuate capital letters in Spanish: *CONFITERIA*. Foix took advantage of both word forms and writing practices in choosing the name of his establishment and in its design. Signs that could be read in the cakeshop were also Cat./Sp. *ESPECIERIA MODERNA* ‘Modern Spicery’ or Cat./Sp. *postres del país* ‘desserts of the country’, among others. Foix’s use of strategic bivalency was not a claim to “neutrality”: it was, in fact, an “act of identity”.

5 Socio-cultural consequences

5.1 Uneven sociolinguistic profiles in the Catalan-speaking area

Fishman (1991, 287–324) mentions the recovery of Catalan and its current status as one case where “reversing language shift” (*RLS*) has proven successful – in terms of language use, language functions and language rights –, side by side with Modern Hebrew and Québécois French. The assertion is partly justified if one brings into focus the “Autonomous Community of Catalonia”, which is an administrative region in Spain – not at all coextensive with Catalonia, neither conceptually nor even territorially –, and leave the rest of the Catalan-speaking area in the shadows.

The Catalan-speaking area is anything but homogeneous. In some territories language loyalty is high and the legal status of the language and the recognised rights of its speakers are well established.²⁸ In other territories the language and its speakers are in a precarious situation or even immersed in language shift or close to language death. Catalonia, Northern Catalonia in France, Valencia, the Balearic Islands, the Aragonese Strip (*la Franja*), *el Carxe* – a handful of small villages in Murcia –, Andorra and the town of *l’Alguer* in Sardinia are uneven settings in uneven contexts – to say nothing about the Catalan gypsy diaspora in France or Catalan worldwide (↗25 Catalan Worldwide).

²⁷ Sarrià, an old village close to Barcelona, is now a borough of the city. It was annexed in 1921.

²⁸ Regrettably, often subject to political ups and downs.

5.2 Language maintenance and shift in the Catalan Countries

Language shift is the process by which an entire people or speech community – or a significant portion of it – abandons their traditional language and comes to use another one that has to an extent been previously learned.²⁹ The former is the receding language and the latter is the dominant language. In this process the receding language loses domains, functions, verbal resources and speakers, and more often than not territory.

Language shift in the Valencian Country was rooted in earlier centuries. It started in the 16th and 17th centuries with the bilingualisation of the aristocracy, who later dropped the local language. In the mid-18th century, an incipient innovative bourgeoisie maintained Catalan at home, but language shift continued, and it expanded downwards among the urban middle class in the 19th century (Ninyoles 1969). In the mid-20th century language disruption became widespread, especially in urban areas (Montoya 1996, 40).

Although the conditions for language shift have been slowly established over centuries in European territorial minorities in general, the process itself was greatly fuelled by the breakdown of the Ancien Régime and the emergence of a world order based on modern nation-states.

A key period in which language shift gained momentum was the industrial revolution in Europe, which entailed a massive migration from the countryside to urban centres.

For many of these minorities, a turning point in the process was in the aftermath of the Second World War by the mid-20th century. This can be observed in Northern Catalonia (Peytaví 2016, 28–29) and in l'Alguer (Chessa 2011, 207), as well as in the Valencian Country (Montoya 1996, 37–40).

Far from creating an irreversible breach, the prolegomena of the industrial revolution in Catalonia provided the basis for strengthening links between the countryside and the developing coastal regions, thanks to the institution of the *hereu*, 'sole heir', rooted in medieval origins. This institution granted the family assets to the eldest son, or heir, avoiding the division of family assets in an economy based on agriculture. This system entailed a series of rights and obligations between the *hereu* and their younger siblings, who usually moved away to earn a living, often with the *hereu*'s support, providing a steady stream of artisans, traders, merchants, foremen, factory managers and workers to the urban centres. Industrial revolution and rural society – *la fàbrica*, 'the factory', and *la masia*, 'the farmhouse' – maintained strong ties, which contributed to the balanced development of the country (Vilar 1962, vol. 3, 567–568,

²⁹ Actually, language shift – usually referred to as a social process – may refer to the linguistic biography of an individual too.

580–581). Social networks were not cut off, but rather contributed to the maintenance of the traditional language.

This notwithstanding, a functional differentiation between both languages still persisted. Spanish was the written language – as well as being used by newspapers and all kinds of books or grey literature, it was the language of bills, administrative correspondence, commercial letters, financial books, advertising and signs, even if business was run and deals were closed in Catalan. This was a clear symptom of a diglossic situation, motivated by the exclusion of Catalan from school and public life.

In spite of the attempt by certain urban bourgeois families to address their children in Spanish around the middle of the 19th century and in spite of some defections after the victory of the fascists in the Spanish Civil War (1936–1939), language shift did not significantly impinge on Catalan society. Nonetheless, Catalonia was about to suffer a dramatic change in its demographic structure resulting from migratory waves from economically deprived areas of Spain during the 20th century and from overseas in the first decades of the 21st century (↗18 Language Demography). This change may have contributed to the minoritisation of Catalan in its stronghold, but not strictly because of the disruption of language transmission by native speakers.

Disruption of language transmission in the Balearic Islands started over the last few decades under the pressure of increasing immigration and residential tourism, whereas in the Aragonese strip (*la Franja*) – a traditional stronghold where Catalan was universally understood and used (Querol et al. 2007) – it seems that the maintenance of intergenerational transmission may collide with occasional disruption from family members of the same generation – spouses, siblings –, due to recent demographic minoritisation, involving mixed marriages (Sorolla 2015, 328–334; 2017). Andorra, an independent state in the Pyrenees with Catalan as its only official language, has not experienced language shift (Querol et al. 2007) – though migratory pressure is keenly felt there too.

Current regional legislation protects Catalan in the Spanish areas, with varying aims, efficacy and results in each, since the extent to which legislation is implemented does not always entail reversing language shift (*RLS*) (↗19 Language Law and Language Policies; ↗20 Teaching and Learning of Catalan).

Northern Catalonia and l'Alguer are the weakest links in the chain of territorial language maintenance (Querol et al. 2007; Chessa 2007). The former has suffered attempts to efface Catalan since well before the French Revolution (Iglesias 1998); the process intensified under the Republic's policies intended to eradicate regional languages by repressive means, through the spread of universal education – which facilitated people's access to French – and as an aftermath of mobilisation during both World Wars. Revitalisation efforts have been launched in a hostile environment – 1976 saw the establishment of the first *bressola* ('cradle') or (private) school to teach in Catalan.

The endurance of Catalan in l'Alguer from medieval times up to the 21st century in otherwise alloglot surroundings is noteworthy. The main (if not the sole) language

of communication among members of the Catalan linguistic community must still have been, at least until the end of the first half of 20th century, the local variety of Catalan or *Alguerese* (Chessa 2011, 104). Census data show that, in 1921, virtually all residents of *l'Alguer* (99.64 %) could be considered Catalan-speakers (Pueyo 1996, 77–78; Grossmann 1983, 13). Local Catalan, however, has been slowly dying out as the result of a series of overlapping socioeconomic, sociocultural and political factors in the 1950s and 1960s, including the opening of modern factories in the area and the development of mass tourism, which led to important changes and opportunities in people's lifestyles and a decline in traditional socioeconomic activities;³⁰ the expansion of national ideology and discourse – contributing to emerging social stigma –; the diffusion of the national language – leading to the stigmatisation of the local language –; the loosening of social networks inasmuch as the town centre was abandoned by locals – breaking off frequent interaction with each other –; the loss of parents' language abilities in the receding language – making transmission harder. There have been bottom-up efforts towards revitalisation, and also top-down legislation passed by the Autonomous Region of Sardinia (1997) and the Italian Republic (1999), without consequences.

Neither the French Republic nor the Italian Republic has yet ratified the European Charter for Regional or Minority Languages (adopted by the Council of Europe on 1992).

Key moments in language shift happen when the dominant language pervades the homes of receding language speakers, eventually leading to the disruption of language transmission (Mas/Montoya 2011a; 2011b). The case of a dominant language forcefully enroaching upon a receding language speakers' home through secondary socialisation, the media, and other factors, must be distinguished from the case of linguistically mixed marriages. Obviously, there is not one single pattern that all mixed marriages opt for, as shown in the collection of Boix-Fuster/Torrens (2011). Nonetheless, a key turning point is generally the birth of the first child. At this point, a decision is usually taken, either consciously or not, explicitly negotiated or not, over which language to use. In cases in which parents used to talk to each other in one of the languages present, this decision is simple. Otherwise, the choice does not necessarily favour one language, but language mixing may emerge (Boix-Fuster/Torrens 2011, 124–126) – whatever attitude parents take towards “pure” or “mixed” speech forms.

30 Crafts and guilds are domains that favour language maintenance and the retention of large amounts of traditional terminology (Simon 2013).

5.3 An ethnographic approach: the voice of the speakers

Language shift has been approached using sociological and ethnographic methods.

The former methods rely heavily on a series of censuses or surveys in order to access relevant quantitative data and apply quantitative analyses. Explanations, however, cannot be mechanistic; indeed, survey data require interpretation. Thus, the progressive decrease in the number of Catalan speakers in Catalonia has not been so much the result of language shift as of demographic change in the 20th century, whereas Valencia was strongly affected by shift.

The latter methods rely on participant observation, recording in-depth interviews or group debates as well as natural conversation or performances, and on other means of accessing relevant qualitative natural data. Ethnographic methods always give voice to the involved speakers, who try to make sense of their current experience. The researcher tries to find out the rationale that lies behind this and develops an interpretative narrative of facts.

The speakers' voice may sound similar in different settings or viceversa. Let us hear how differently an old man in l'Alguer and an old Catalan gypsy in Perpinyà experience and express their feelings.

The first speaker was narrating his linguistic biography.³¹ Although he was born to an Alguerese/Sardinian family, he received neither one language nor the other, but Italian. L'Alguer was about to undergo language shift at that time –but that man belongs to a generational group who learnt local Catalan from their playmates outside of the home in natural settings and used mainly local Catalan. This situation is nostalgically compared with the one the children in l'Alguer confront nowadays:

Quan tenia de jugar amb els altres era obligatori en aquell temps saber parlar l'alguerès, perquè no hi havia altra possibilitat: o eres... o parlaves l'alguerès o eres una mica emarginat. [...] he conegut una Alguer diferent i avui veig que aqueixa Alguer no hi és més. Avui els minyons no són com era jo, no tenen necessitat d'emparar l'alguerès per jugar al carrer, anzi, si no parlen l'italià no juguen. La substitució lingüística en aquí és ormai feta, no hi ha estat ninguna reacció social contra aqueixa substitució lingüística [...] (SPEAKER_014_M62).

'[...] when I had to play with the other children, it was compelling, at that time, to know and speak Alguerese, because there was not... no alternative: either you were..., either you spoke Alguerese or you were somehow excluded, [...] I knew a different Alguer, and nowadays I see that this Alguer is... no longer exists. Today children are not like my... like I was. They don't have the need to... learn Alguerese to play in the street, on the contrary... if they don't speak Italian, they don't play. This is... Language shift is already completed here. There has been no social react- reaction against this language shift [...].'

31 For a research making crucial use of the method of linguistic biographies in l'Alguer, see Simon (2018).

A somehow opposite reaction facing language shift shows up in the following extract from an interview between the researcher and a Catalan gypsy (J. S.) in the south of France, namely the latter's answer to a question posed by the former.

–Tant com els vells seran aquí irà bé. Mes durarà pas temps. De la manera que va el poble gitano és una catastrophe. El barregen amb la llengua d'on se trapen [es troben]. La meua noia i el meu noi de tant en tant, tenen una escapada en francès, i jo dic: Ei! Poc a poc, sóc aquí encara! El dia que hi serem pas més, feu lo que voldreu (Escudero 2014, 214).

–As long as as the old people are here [= as long as as they are alive] it will go well. But it won't last long. The way the gypsy people are going is a catastrophe. They mix it [Catalan] together with the language spoken where they are. From time to time, my daughter and my son break into French, and I say: Hey! Bit by bit, I am still here! The day I'm gone, you can do as you like.

As is the case with Catalan in l'Alguer, Gypsy Catalan suffers under the overwhelming pressure from the national language, French in this case, and language reproduction is not granted. However, speaker J. S. is proud of his received though receding language, he is reluctant to use the dominant language and, although he is conscious that Gypsy Catalan is not at its best, demonstrates that he still wants to maintain control over the fate of his language and over his descendants' language usage.

Furthermore, we can ascertain only from shifting speakers themselves or from their children or grandchildren how they try to make sense of their option not to transmit Catalan, how they rationalise it or how they construe both its causes and their responsibility in the process. A typology of such rationalisations is sketched out by Montoya (1996, 173–176), in line with Querol (1989).

While the shifting generation causes the disruption of language transmission after not having passed on the language to their offspring, occasionally the elder generation, the grandparents, becomes a source of knowledge (Argenter 2013). Dorian (1981, 107–108) reports that children who maintain close contact and emotional involvement with a (female) member of the second ascending generation develop positive values towards the traditional language and are eager to learn it. She terms this the “grandmother factor” (Dorian 2014, 87).

The following interview fragment from an Alguerese speaker bears witness to how language shift entailed the loss of a significant verbal genre in traditional culture, i.e., prayers – a genre that, so it seems, was rather gendered and aged locally, mainly associated with grandmothers. Asked whether she used to say prayers in Alguerese when a child, a woman in her seventies says:

–No. A mi no me n'han dit mai. Sempre en italià. Invece hi era la iaia que les dieva, ma a mi me són esfugides aqueixes preguieres. Les sap un amic meu. M'ha dit: “jo l'he cultivada sempre, que ma iaia, petit, era sempre en casa, i cada nit – també grans – mos dieva: dieu-vos la preguiera no sep com la...”. Me l'ha dit també l'altra volta que só anada a me fer els cabells. I m'ha dit que se la diu cada nit ell. “Me la dic en alguerès com me l'ha emparada la iaia”. Invece de diure l'Ave Maria ... ecco. No. Se la diu en alguerès. I és bella... (Argenter 2013, 153–154).

–‘No. I’ve never been told. Always in Italian. Instead, there was my grandma who said them, but these ... these prayers fled away from my heart. A friend of mine knows them. He told me: “I always cultivated it, ’cause when I was a child my grandma always stayed at home, and every night she said – also to adults –: say your prayer. I don’t know how...”. He retold it again another time when I went to the hairdresser. And he told me that he says the prayer every night, he... “I said my prayer in Algerese as I learned it from my grandma”. Instead of saying the Ave Maria...that’s it. No, he says it in Algerese. And how beautiful it is...”.’

5.4 Language attrition and obsolescence

The “language shift” paradigm (e.g. Fishman 1964; Ninyoles 1969; Aracil 1982) was defined in terms of decreasing language function, a progressive restriction of domains. The receding language is first abandoned in formal and public domains up to the point that it is restricted to the local neighbourhood and finally to the home, until the language ceases to be used or transmitted.

The “language death” paradigm (Dorian 1981; 1989) addresses “terminal communities” – in which shift is about to be completed, and there is no more language reproduction –, and focuses on the progressive reduction of language form. While elder generations are still fluent speakers, younger generations have a poorer command of the receding language and the youngest – “semispeakers” (Dorian 1981, 107; 2014, esp. part 2) – are no longer fluent in the language, though they may retain some communicative competence. The more the speakers need the dominant language to express things, the less they are able to say things in the receding language. In the process, the grammar shrinks – in other words, speakers cast the language into oblivion. Along the way they suffer language attrition while the language fades into obsolescence (Dorian 1989).

Montoya (2000a; 2000b) studied the phenomenon among terminal speakers in Alacant and Montoya/Jofre (2002) studied it in Argentinian villages among generations of descendents of Majorcan migrants who left in the mid-19th and early 20th centuries (Montoya 2017).

Compressed and unbalanced changes affect expressive resources, vocabulary and all levels of structure, always in the direction of simplification and levelling – often implying the loss of marked forms. There is a lot of interference, but most changes cannot be analysed as interferences from the dominant language, but as structural losses in the receding language resulting from language attrition.³² Let us mention the following:

Generalized regularisation and reduction of inflectional paradigms:

³² In the usual case, interference demands an interfering element and an interfered element.

Clitics:

Article:

les tinc aquí, es nores (Montoya 2000a, 70; 2000b, 100).

'They are here with me [I have them here], my [the] daughters-in-law.'

The masc. vs. fem. forms of the Catalan article, *el/la*, are retained in singular, but in Alacantine Catalan the distinction is lost in the plural in favour of the new unmarked form *es* (from unmarked masc. *els*).

Pronoun system:

Encara lis parle en valencià (Montoya 2000b, 111).

'She still speaks Valencian to them.'

Lis (instead of *els*) is the analogical plural of *li* – this change is rather widespread in Alacantine. Moreover the form *li* may be heard used as plural among those whose first language is Spanish:

Donà-li consej, sí, però obligà-les a res, no [a les nores] (Montoya 2000b, 112).

*To give advice to them, that's okay, but to force them, not at all [the daughters in law].*³³

Doubling:

astó lo hu va inventar E.S.; era un mecànic molt bo (Montoya 2000a, 70; 2000b, 108).

this it it was invented by E.S.; he was a very experienced mechanic.

The sentence contains a "pleonasm" (*astó...lo hu*) and a doubling codeswitching (*lo hu*: Sp. *lo* + Cat. *ho*). This is an instance of the portmanteau construction in code-switching, which involves the juxtaposition of two synonymous morphemes from two different languages (Chan 2016).

Verbal morphology:

Analogical tendencies are observed in forms such as *escrivisc*, 'I write', – 'incorporating the verb *escriure* in the list of inchoative verbs' –, *dint*, 'saying', gerundival

³³ Standard form of the sentence: *Donar-los consell, sí, però obligar-les a res, no*. Oral form in Central Catalan: *Donà'ls-hi consell, sí, però obligà'ls-e a res, no*.

form of *dir* – ‘tipping this verb into the third conjugation model’ –, *perdet*, ‘lost’ – ‘extending, i.e. regularising, the use of the most common thematic vowel of the second conjugation to the participles of this paradigm, that, nonetheless, are made up with -u- (*perdut*)’ – cannot be attributed to Spanish interference at all, but to Catalan attrition in semispeakers (Montoya 2000a, 71; 2000b, 113–115).

Verbal function:

Certain verbal forms are used in a way that distinctions of mode, tense and aspect are erased (Montoya 2000a, 71). In no case are they attributable to Spanish interference.

Other changes indexing attrition fall under sound patterns, syntax or relexicalisation.

6 Technology and language in mass media: broadening function, shrinking form

The domain of the newspaper press in Catalan broadened with the use of machine-translation: several media issue both a Spanish version and a Catalan version. The presence of Catalan in this domain has increased, and so has language functionality. Since the source language tends to be Spanish with the target language being Catalan, imbalance may arise. This technique then demands an investment in language revisers, since the machines’ output is, after all, still not as perfect as humans’. However, humans are limited and imperfect too. Both imperfections together give rise to incorrect choices, be they lexical, morphological, syntactic or pragmatic. A repeated machine-translation failure is the appearance of words such as Cat. *pressuposat* ‘presupposed’, *mesurades* ‘measured’, *proposada* ‘proposed’, *encarrego* ‘I entrust’, instead of Cat. *pressupost* ‘budget’, *mesures* ‘measures’, *proposta* ‘proposal’, *encàrrec* ‘comission’ (from Sp. *presupuesto*, *medidas*, *propuesta*, *encargo*). These failures are presumably easily detected by the reader.

No more difficult to detect, although more serious – since it affects a core feature of Catalan –, are failures involving the system of verbal clitics. Spanish lacks the equivalents of pronouns *hi* and *en*.³⁴ Many Catalan L2 speakers present an insecure or variable, even random, use of the clitic system – either by omission, excessive use or incorrect choice. The machine does not do any better:

³⁴ And allomorphs (*en*, *ne*, *n*), graphically represented as *en*, *-ne*, *n’*, *’n*, e.g. *en dona* (also dialectal *ne dona*), *donar-ne*, *n’ha donat*, *dona’n*.

[...] *al pla europeu de refugiats, que assignava uns 11.000 a Hongria.* (La Vanguardia, 08/09/2015, 3).

‘[...] against the European plan for refugees, that allocated ca. 11,000 refugees to Hungary.’

Instead of ... *que n’assignava uns 11.000 a Hongria.*

In other cases, where the difference between the correspondences is greater, understanding demands a supplementary effort by means of backward translation. Consider, for instance, the following sentence:

Tres germanes viuen les seves vides independentment fins que el pare malalta (La Vanguardia, 08/09/2015, 14).

‘Three sisters live their lives independently until their sick father.’

The last word is a feminine adjective apparently agreeing with the preceding masculine noun, but the main point is that the syntax is not at all Catalan. The cause of this is that *malalta* translates Sp. *enferma*, and this may be either a feminine adjective or a verbal form (third person, singular, present tense, indicative mood of *enfermar*). The machine fails by “misinterpreting” the verb as an adjective – the corresponding Catalan verb form is *emmalalteix* (from *emmalaltir*, ‘to sicken’). The meaning would be then ‘until their father sickens’.

Next comes a rather clumsy, but multiple-faceted case that bears on lexical false friends, discursive mechanisms, figurative language, and “lost” verbal cultural knowledge:

Mitjançant resolucions clares, que convoquin a cada pal institucional a aguantar la seva espelma (José Maria Brunet, En alerta màxima, La Vanguardia 02/11/2015).

‘By means of clear resolutions, summoning each institutional mast to hold up its candle.’

The writer plays with the Spanish colloquial expression *cada palo que aguante su vela* (cf. Engl. *You’ve made your bed, now lie in it*) – adding a new word: *institucional* and adapting the syntax. The machine erroneously translates Sp. *vela* ‘sail’ as Cat. *espelma* ‘candle’, since this is the meaning of the homograph Sp. *vela*. The human reviser overlooked this error made by the machine. However, two even more relevant points merit attention: the reviser probably ignored the corresponding Catalan fixed expression, that is, *qui sigui frare que prengui candela*, lit. ‘he who is a friar, let him hold the candle’. We know from experience that most Catalans forgot – or never learnt – Catalan proverbs and figurative fixed phrases as they learnt Spanish ones. In fact, the reviser’s main failure was to accept a Spanish idiom as a Catalan one. Moreover, to make it more difficult, the machine yielded an idiomatic Cat. *aguantar l’espelma*, ‘to chaperone’, that has gone unnoticed.

The moral of this story is that, in principle, machine-translation in the media helps broaden the social function of Catalan in this language domain, but eventually

it might – if not correctly supervised – turn mass-media Catalan into a calque language – thus seriously damaging its form.

7 Bibliography

- Aracil, Lluís V. (1977), *Educació i sociolingüística*, Treballs de Sociolingüística Catalana 2, 33–86.
- Aracil, Lluís V. (1982), *Conflicte lingüístic i normalització lingüística a l'Europa nova*, in: Lluís V. Aracil, *Papers de sociolingüística*, Barcelona, La Magrana, 23–38.
- Argenter, Joan A. (2008), *L'Alguer (Alghero), a Catalan Linguistic Enclave in Sardinia*, International Journal of the Sociology of Language 193–194, 28–46.
- Argenter, Joan A. (2013), *Intergenerational Permeability and Cleavages: From Parents as Agents of Language Shift to Grandparents as a Source of Knowledge*. Catalan: a Heritage Language in Alghero, Sardinia, in: Mary Jane Norris et al. (edd.), *Endangered Languages Beyond Boundaries: Community Connections, Collaborative Approaches, and Cross-Disciplinary Research/Langues en péril au-delà des frontières: Connexions communautaires, approches collaboratives, et recherche interdisciplinaire*, Bath, Foundation for Endangered Languages, 150–156.
- Argenter, Joan A./Ferrer, Joan (forthcoming), *Jewish Texts in Medieval Catalan*, in: Guido Mensching/Frank Savelsberg (edd.), *Manual of Judaeo-Romance Linguistics and Philology*, Berlin/Boston, De Gruyter.
- Auer, Peter (1999), *From Code-Switching via Language Mixing to Fused Lects: Toward a Dynamic Typology of Bilingual Speech*, International Journal of Bilingualism 3, 309–332.
- Bakhtin, Mikhail M. (1981), *The Dialogical Imagination*, ed. Michael Holquist, Austin, University of Texas Press.
- Bauman, Richard/Briggs, Charles L. (1990), *Poetics and Performance as Critical Perspectives on Language and Social Life*, Annual Review of Anthropology 19, 59–88.
- Blom, Jan-Petter/Gumperz, John J. (1972), *Social Meaning in Language Structure: Code-Switching in Norway*, in: John J. Gumperz/Dell H. Hymes (edd.), *Directions in Sociolinguistics*, New York, Holt, Rinehart & Winston, 407–434.
- Boix, Emili (1993), *Triar no és trair: identitat i llengua en els joves de Barcelona*, Barcelona, Edicions 62.
- Boix-Fuster, Emili/Torrens, Rosa Maria (edd.) (2011), *Les llengües al sofà. El plurilingüisme familiar als països de llengua catalana*, Lleida, Pagès.
- Casanova, Eugeni (2016), *Els gitanos catalans de França. Llengua, cultura, itineraris de la gran diàspora*, Lleida/Barcelona, Pagès/Càtedra UNESCO de Diversitat Lingüística i Cultural.
- Chan, Brian Hok-Shing (2016), *Portmanteau Constructions, Phrase Structure, and Linearization*, in: Artemis Alexiadou/Terje Lohndal (edd.), *The Grammar of Multilingualism*, Frontiers Research Topics, 95–110. DOI: 10.3389/978-2-88945-012-1.
- Chessa, Enrico (2007), *Enquesta sobre els usos lingüístics a l'Alguer 2004: Llengua i societat a l'Alguer en els inicis del segle XXI*, Barcelona, Generalitat de Catalunya.
- Chessa, Enrico (2011), *Another Case of Language Death? The Intergenerational Transmission of Catalan in Alghero*, PhD thesis, University of London.
- Dorian, Nancy C. (1981), *Language Death. The Life Cycle of a Scottish Gaelic Dialect*, Philadelphia, University of Pennsylvania Press.
- Dorian, Nancy C. (ed.) (1989), *Investigating Obsolescence. Studies in Language Contraction and Death*, Cambridge, Cambridge University Press.
- Dorian, Nancy C. (2014), *Small-Language Fates and Prospects. Lessons of Persistence and Change from Endangered Languages. Collected Essays*, Leiden/Boston, Brill.

- Escudero, Jean-Paul (2014), *Enclavaments lingüístics dels gitanos catalanoparlants al Rosselló i a Occitània: un estat de la qüestió*, in: Joan A. Argenter (ed.), *Enclavaments lingüístics i comunitats locals. El català a Sardenya, al Carxe i entre els gitanos catalans de França*, Barcelona, Institut d'Estudis Catalans, 209–219.
- Fishman, Joshua (1964), *Language Maintenance and Language Shift as a Field of Inquiry*, *Linguistics* 9, 32–70.
- Fishman, Joshua A. (1965), *Who Speaks What Language to Whom and When?*, *La Linguistique* 1, 67–88.
- Fishman, Joshua A. (1972), *Domains and the Relationship Between Micro- and Macrosociolinguistics*, in: John J. Gumperz/Dell H. Hymes (edd.), *Directions in Sociolinguistics*, New York, Holt, Rinehart & Winston, 435–453.
- Fishman, Joshua A. (1991), *Reversing Language Shift: Theoretical and, Empirical Foundations of Assistance to Threatened Languages*, Bristol, Multilingual Matters.
- Gal, Susan (1979), *Language shift. Social Determinants of Linguistic Change in Bilingual Austria*, New York/San Francisco/London, Academic Press.
- Goffman, Erving (1981), *Forms of Talk*, Philadelphia, University of Pennsylvania Press.
- Grossmann, Maria (1983), *Com es parla a l'Alguer?*, Barcelona, Barcino.
- Gumperz, John J. (1982), *Discourse Strategies*, Cambridge, Cambridge University Press.
- Hill, Jane H. (1993), *Structure and Practice in Language Shift*, in: Kenneth C. Hyldenstam/Åke Viberg (edd.), *Progression and Regression in Language*, Cambridge, Cambridge University Press, 68–93.
- Hymes, Dell H. (1972), *Models of the Interaction of Language and Social Life*, in: John J. Gumperz/Dell H. Hymes (edd.), *Directions in Sociolinguistics*, New York, Holt, Rinehart & Winston, 35–71.
- Iglesias, Narcís (1998), *La llengua al Rosselló, qüestió d'Estat. La integració lingüística del Rosselló a França (1659–1789)*, Vic, Eumo.
- Jakobson, Roman/Halle, Morris (1956), *Fundamentals of Language*, The Hague, Mouton.
- Kailuweit, Rolf (2002), *Català heavy – català light: una problemàtica de la “lingüística de profans”*, *Zeitschrift für Katalanistik* 15, 169–182.
- Mas, Antoni/Montoya, Brauli (2011a), *La transmissió lingüística del català: estat de la qüestió i avaluació analítica*, *Treballs de Sociolingüística Catalana* 21, 95–103.
- Mas, Antoni/Montoya, Brauli (2011b), *La transmissió familiar del valencià*, València, Acadèmia Valenciana de la Llengua.
- Montoya, Brauli (1996), *Alacant: la llengua interrompuda*, València, Denes.
- Montoya, Brauli (2000a), *L'atròfia lingüística del català a la ciutat d'Alacant*, in: *Jornades de la Secció Filològica de l'Institut d'Estudis Catalans a Elx i a la Universitat d'Alacant (16 i 17 d'octubre de 1998)*, Barcelona, Institut d'Estudis Catalans, 65–74.
- Montoya, Brauli (2000b), *Els alacantins catalanoparlants: una generació interrompuda*, Barcelona, Institut d'Estudis Catalans.
- Montoya, Brauli (2017), *La llengua dels mallorquins en San Pedro (Argentina)*, Barcelona/Palma, Publicacions de l'Abadia de Montserrat/Universitat de les Illes Balears.
- Montoya, Brauli/Jofre, Ana (2002), *La atrición en las lenguas inmigradas. El caso del catalán de los mallorquines en la Argentina*, in: *Actas do i Simposio Internacional sobre o Bilingüismo*, Vigo, Universidade de Vigo, 778–789.
- Muysken, Pieter (1997), *Code-Switching Processes: Alternation, Insertion, Congruent Lexicalization*, in: Martin Pütz (ed.), *Language Choices: Conditions, Constraints, and Consequences*, Amsterdam/Philadelphia, Benjamins, 361–380.
- Ninyoles, Rafael Lluís (1969), *Conflicte lingüístic valencià*, València, Edicions 3 i 4.
- Pericay, Xavier/Toutain, Ferran (1986), *Verinosa Llengua*, Barcelona, Empúries.
- Peytaví, Joan (2016), *El català al nord de Catalunya a principi del segle XXI*, Barcelona, Institut d'Estudis Catalans.

- Pueyo, Miquel (1996), *Tres escoles per als catalans*, Lleida, Pagès.
- Pujolar, Joan (2001), *Gender, Heteroglossia and Power: A Sociolinguistic Study of Youth Culture*, Berlin/New York, de Gruyter.
- Querol, Ernest (1989), *El procés de substitució lingüística. La comarca dels Ports com a exemple*, in: *Miscel·lània 89*, València, Generalitat Valenciana, 85–196.
- Querol, Ernest, et al. (2007), *Llengua i societat als territoris de parla catalana a l'inici del segle XXI: L'Alguer, Andorra, Catalunya, Catalunya Nord, La Franja, Illes Balears, i Comunitat Valenciana*, Barcelona, Generalitat de Catalunya.
- Rampton, Ben (1998), *Language Crossing and the Redefinition of Reality: Expanding the Agenda of Research on Code-switching*, in: Peter Auer (ed.), *Code-switching in Conversation: Language, Interaction and Identity*, London, Routledge, 290–317.
- Silverstein, Michael (1992), *Metapragmatic Discourse and Metapragmatic Function*, in: John Lucy (ed.), *Reflexive Language: Reported Speech and Metapragmatics*, Cambridge, Cambridge University Press, 33–58.
- Silverstein, Michael/Urban, Greg (1996), *The Natural History of Discourse*, in: Michael Silverstein/Greg Urban, *Natural Histories of Discourse*, Chicago, University of Chicago Press, 1–17.
- Simon, Sophia (2013), “Parole che dicono tutta una frase”. *Fachtermini im Handwerk und sektorenspezifische Kommunikation als Domänen der Spracherhaltung in L'Alguer (Sardinien)*, in: Ulrich Hoinkes (ed.), *Die kleineren Sprachen in der Romania*, Frankfurt am Main, Lang, 207–234.
- Simon, Sophia (2018), *L'identità lingüística e culturale degli abitanti di Alghero (Sardegna) – un approccio biografico-linguistico*, Wien, Praesens.
- Sorolla, Natxo (2015), *Tria de llengües i rols sociolinguistics a la Franja des de la perspectiva de les xarxes socials*, PhD thesis, Universitat de Barcelona.
- Sorolla, Natxo (coord.) (2017), *Els usos lingüistics a la Franja*, Barcelona, Generalitat de Catalunya, Departament de Cultura.
- Thomason, Sarah G./Kaufmann, Terrence (1988), *Language Contact, Creolization, and Genetic Linguistics*, Berkeley/Los Angeles/London, University of California Press.
- Vilar, Pierre (1962), *La Catalogne dans l'Espagne moderne*, 3 vol., Paris, S.E.V.P.E.N. (Bibliothèque générale de l'École Pratique des Hautes Études. VIe Section).
- Weinreich, Uriel (1970, 1953), *Languages in Contact. Findings and Problems*, The Hague, Mouton.
- Woolard, Kathryn A. (1989), *Double Talk: Bilingualism and the Politics of Ethnicity in Catalonia*, Stanford, Stanford University Press.
- Woolard, Kathryn A. (1999), *Simultaneity and Bivalency as Strategies in Bilingualism*, *Journal of Linguistic Anthropology* 8, 3–29.
- Woolard, Kathryn A. (2016), *Singular and Plural. Ideologies of Linguistic Authority in 21st Century Catalonia*, Oxford, Oxford University Press.

F. Xavier Vila

18 Language Demography

Abstract: This chapter starts with a review of some historical traits that have characterized the Catalan language community from a demolinguistic point of view, showing that the history of the language does not follow the commonly acknowledged pattern: a medieval golden age, followed by decadence in the Modern Age, and a subsequent renaissance in the wake of Romanticism. Data suggest, on the contrary, that the most essential demolinguistic transformations affecting Catalan have taken place during the last two centuries. The article then turns to the current situation of Catalan from a variety of perspectives: in terms of language proficiency, as a first language, and in terms of identification and language use. The data highlight the considerable demolinguistic differences existing among territories, regions and even cities, and shed light on the role played by a number of social variables such as family origin, social class, sex and age.

Keywords: language demography, Catalan, language proficiency, language use, inter-generational transmission

1 A demolinguistic approach to Catalan language history

The demography of language or *demolinguistics* is the statistical study of populations and their flows from the point of view of the language(s) they know and use. The demolinguistic reality of any language is deeply intertwined with political, economic, cultural, literary and other social phenomena but, especially in multilingual contexts, the evolution of each of these variables often exhibits significant differences. Therefore, if the vitality and evolution of any given language is to be fully understood, all of these variables deserve to be considered in their own right.

This chapter focuses on the demolinguistic evolution and current situation of Catalan,¹ which is a good example of how political, cultural and demolinguistic factors

1 A relatively abundant demolinguistic literature has been produced in Catalonia over the last three to four decades: see <http://llengua.gencat.cat/ca/el-catala/coneixement-i-us/> or <http://www.ub.edu/cusc/demolinguistica/>. Especially relevant for this article are two sets of surveys that were conducted in a rather coordinated way in all Catalan language territories, here referred to as EUL 2003–2004 (*Enquesta d'Usos Lingüístics*, i.e. 'Survey on Language Uses', in 2003–2004; see Querol et al. 2007) and EUL 2013–2015 (Direcció General de Política Lingüística 2016; 2018). First data from Catalonia's EUL 2018 were about to be published at the time of editing this chapter.

are interconnected but should not be confused. The history of Catalan has often been explained based on the literary and political events that have punctuated its existence. According to this narrative (↗11 The Growth and Expansion of Catalan (1213–1516) through to ↗14.2 From Pompeu Fabra to the Present Day: Language Change, Hindrance to Corpus and Status Planning), Catalan was born in the Middle Ages, rose to its zenith with the expansion of the Crown of Aragon, entered literary and even linguistic *Decadence* after the union with Castile, and began a period of renaissance in the 19th century against a backdrop of Romanticism and modern nationalism. Irrespective of the exactitude of this discourse in political and literary terms, in the next sections we will see that the demolinguistic evolution of the language does not coincide with this interpretation, and that a demolinguistic approach can improve our understanding of both the past and the current situation of Catalan.

2 The demolinguistics of Catalan up to the 19th century

Catalan is what it is today because a number of peoples – and their languages – came into contact, traded, fought and eventually intermingled. The definition of Catalan's historical homeland is the result of the sociopolitical and demographic process that elapsed between the creation of the Catalan counties in the 9th century and the conquest and colonization of the Balearic Islands and the Kingdom of Valencia in the 13th century, with the addition of l'Alguer in the 14th century (↗11 The Growth and Expansion of Catalan (1213–1516)). Although Latin was still the language of Christian liturgy and the main written language, the majority of the population of all social classes in these territories – with the exception of Arabic-speaking minorities – spoke nothing but Catalan. As a consequence, between the 13th and 15th centuries the language progressively occupied more and more formal domains as a language of public administration, literature and even science, both orally and in written form.

The political union with the Kingdom of Castile undermined the position of Catalan as a formal language, especially in the 16th and 17th centuries, when the Crown of Aragon became a peripheral dominion under the Habsburgs (↗12 The Origins of Modern Catalan: Cultural and Linguistic Evolution). The marriage of Ferdinand and Isabella paved the way for a number of Catalan aristocratic families such as the Requesens or the Montcada to arrange marriages with their Castilian peers and, once they had become established around the new royal court in Castile, their descendants gave up Catalan as their family language. The new world superpower created by the Habsburgs also attracted some members of the Catalan nobility to serve as imperial servants or in its armies. All of these changes, though, only concerned a handful of individuals that became expatriates and, even if they retained their properties at home, had very little impact in the Catalan-speaking territories themselves.

More significant in sociolinguistic terms was the arrival of a certain number of royal delegates who served in the royal institutions, the Inquisition, the secular and regular church, the army, etc., and used Castilian as the language of the central institutions to address the local elites. Their position encouraged familiarity with this language among the upper classes and in highly educated circles from the 16th and 17th centuries onwards and fueled its prestige as the main language of culture (712 *The Origins of Modern Catalan: Cultural and Linguistic Evolution*). The practice of Catalan as a language of culture and formal usage was further debilitated by the annexation of Roussillon by France in 1659 and by the annexation of the Catalan-speaking countries to Castile after 1714. The Bourbon monarchs exacerbated the hegemony of the state language both in Northern Catalonia, where French has been the official language since 1700, and in Spain, where Madrid's language became the sole official language after the War of Succession. As in many other European societies during the *Ancien Régime*, the use of the language of the monarchy functioned in the Catalan Countries as a social marker that distinguished the aristocrats from the commoner in two senses (Marfany 2001): on the one hand, it was the language to be used for the most distinguished occasions; on the other, its learning was restricted to the ruling classes.

But this *aristocratic diglossia*, as we could name it, does not mean that Catalan nobles abandoned the use and intergenerational transmission of Catalan, and there is evidence that their everyday conversations continued to take place in their first language (Marfany 2001). The main exception to this rule is that of Valencia, where the castilianization of the Court after the Germanies War led to the abandonment of Catalan by the nobles (Ninyoles 1969). As far as the rest of society is concerned, the majority of Catalan speakers remained basically monolingual and had only minor contact with the language of the royal institutions. In other words, neither the constitution of the Habsburg Empire nor the establishment of the absolutist Bourbon monarchy significantly modified the demolinguistic structure of the Catalan societies. Between the 16th and 19th centuries, the Catalan-speaking territories remained ethnolinguistically rather homogeneous. In fact, the main demolinguistic change during this period was the expulsion in 1609 of the *moriscos*, i.e. the Arabic-speaking descendants of the inhabitants of Muslim territories that had come under Catalan control in the Middle Ages. This expulsion, which entailed the sudden banishment of around one third of the Valencian population, implied the loss of the only significant (in demolinguistic terms) ethnic minority in the Catalan-speaking territories (Montoya Abad 2009) and, therefore, an *increase* of the territories' ethnolinguistic homogeneity.

Homogeneity is, of course, a relative term. The Catalan Countries have been the natural corridor between Europe and Africa for millennia, and the bigger cities such as Barcelona and Valencia were especially accustomed to the presence of traders, troops, visitors, sailors, pirates and slaves from a variety of origins. In other words, the presence of foreigners has never been an anomaly in these territories, as witnessed by the multitude of Catalan family names corresponding to nationalities such as Alemany ('German'), Danès, ('Danish'), Gascó ('Gascon'), Llo(m)part ('Lombard'),

Navarro ('Navarrese') or many others. But although some of these population movements reached a significant demographic dimension, especially the immigration of Occitan speakers to some regions in Catalonia in the 16th and 17th centuries (Nadal/Giralt 1960), these immigrants totally assimilated into Catalan culture and left very few demolinguiistic traces behind. Contrary to what happened in Central and Eastern Europe, where a myriad of language islands bear witness to historical migratory processes, in the Catalan-speaking territories until the 20th century, immigration never really led to the formation of separate linguistic communities. Even gypsies, who arrived in the 15th century and have retained a strong sense of being a distinct community to the present day, adopted Catalan as their only language of communication between the 18th and 19th centuries (Escudero 2000).

The main exception to this demolinguiistic rule is to be found in the southernmost region of the linguistic area. Catalan was introduced in the Kingdom of Murcia in 1266 when the Catalan King James I repopulated this territory with Catalans after quashing the Muslim revolt in a region that had been conquered by Castilians only two decades previously. Due to continuous Castilian immigration, Catalan was progressively superseded and eventually disappeared from Murcia in the 16th century. This process of castilianization in the south spilled over the political borders and spread to the neighboring Valencian *comarca* of Vega Baja/Baix Segura, which castilianized in the 18th and 19th centuries (Montoya Abad 1986). In the 20th century, the city of Alacant and surrounding areas experienced severe, although not complete language shift (Montoya Abad 1996).

3 Demolinguiistic changes in the 19th and 20th centuries

The traditional framework of demolinguiistic stability and relative ethnolinguiistic homogeneity started to change significantly in the 19th century, even more dramatically in the 20th century, and has continued to evolve in the 21st century. The major sociodemographic changes started with the fall of the Ancien Régime, which opened the door to a more meritocratic society and facilitated social mobility. The industrialization of certain regions started to attract growing numbers of rural migrants from more and more distant places, a phenomenon encouraged by the dramatic reduction in the fertility rate of Catalan women since the 19th century (Cabré Pla 1999). These socioeconomic changes were accompanied by three different sociolinguiistic phenomena with a distinct demolinguiistic dimension: first, the spread of proficiency in the state language; second, a massive language shift towards the state language in some regions; and third, mass immigration of non-Catalan speakers.

The first linguistic correlate of the aforementioned changes was the spread of competence in the official state language (↗20 Teaching and Learning of Catalan).

Little by little, all over the Western world, the ruling elites espoused a new ideology which actively sought to establish the “national language” as the sole language of culture for all social classes, not only for the elites (Anguera 1997; Marfany 2001). Thus, the authorities in Spain, France and Italy progressively set up free, universal education in the designated *national* languages (Pueyo 1996). It should not be forgotten that, in comparison with the Ancien Régime, where status was acquired by birth, in industrial societies proficiency in official languages became a commodity that helped social mobility (Bourdieu 1982; Heller/Duchêne 2012). The combination of public policies with social incentives and new ideologies effectively led to a progressive increase in proficiency in the state languages in all Catalan-speaking territories. The unfortunate lack of linguistic censuses and statistics makes it very difficult to discern the exact chronology of this bilingualization, but there is plenty of data to suggest that, in general terms, receptive abilities in Castilian, French or Italian only became common – although by no means universal – from the mid-19th century onwards. As far as active proficiency in the state language is concerned, this only became really universal in the 20th century (Anguera 1997; Marfany 2001). Have in mind that the literacy rate in Spain had only reached 50 % among the young population by around 1920 (de Gabriel 1997, 220). As late as 1931, Fernando Varela, a socialist MP stated at the Spanish Cortes:

[E]n una gran extensión de Cataluña –yo he vivido mucho en ella y conozco esto bien por experiencia– no hay nadie que sepa hablar castellano, no se habla más que catalán

The spread of proficiency in the state language accelerated during the first half of the 20th century and became universal during its second half, hand in hand with the progressive extinction of monolingual Catalan speakers due to either informal language learning on the streets, in military service, via the new mass media, all in the state languages, or to the passing of older speakers.

The second demolinguistic change that took place over the last two centuries was the massive language shift towards the state language in some areas. Whereas such a language shift was rare in the Ancien Régime, the new dynamics of social mobility beginning during the 19th century shook traditional sociolinguistic balances all over Europe. The very idea of “*the* language of the nation” and the rewards it carried with it in terms of social mobility were an invitation to transcend diglossia and dive head-first into total language shift, including the interruption of intergenerational language transmission, a process traditionally described in Catalan sociolinguistics as *language conflict* (Ninyoles 1969; Vila 2014). Between 1830–1850, some bourgeois families started to address their children in Castilian in Valencia and Barcelona, and over the following decades the process spread to other big cities such as Alacant (Anguera 1997; Montoya Abad/Mas i Miralles 2011). But the evolution of this language shift was to be very different in each territory. In Catalonia, the Catalanist mobilization – which, in practical terms, proposed that Catalan was the authentic national language of Catalans – stopped the process of language shift (Anguera 1997) and, in spite of some

defections, especially after Franco's victory (1939), the abandonment of Catalan as the language in the home never reached significant numbers beyond the upper class (McRoberts 2001). In Northern Catalonia and in l'Alguer, on the contrary, the process intensified and reached its peak between the 1940s and 1960s, when middle and working classes stopped teaching Catalan to their children (Bernardó/Rieu 1976; Chessa 2011). In the Valencian Country, the language shift affected the three biggest urban areas of Alacant-Elx, Valencia and, a little bit later, Castelló de la Plana, but did not spread to smaller cities or rural areas. In the Balearic Islands, the interruption of language transmission only seems to have started over the last few decades (Fabà/Montoya 2012), whereas in la Franja it seems to be incipient (Sorolla 2017b). Andorra, an independent state with Catalan as its only official language has never experienced language shift towards any foreign language (Querol et al. 2007).

Finally, during the 20th century, the Catalan language territories experienced a third fundamental demolinguiistic change: the massive immigration of non-Catalan speakers due to socioeconomic changes which led to the development of primary industry and, later on, the service economy. Again, this evolution was different in each territory.

Catalonia was the first territory to experience massive immigration of non-Catalan speakers. During the first decades of the 20th century, the need for workers in the flourishing industries around Barcelona started attracting people from beyond the borders of the Catalan language. Within a couple of decades, thousands of people from Murcia and Eastern Andalusia had moved to Barcelona and other industrial areas in the vicinity (Domingo 2014b). Though the 1929 crisis stopped this influx, if only temporarily, in 1930, 14.5 % of the population living in the province of Barcelona were immigrants (Espinet 2001, 107–108).

The Spanish Civil War (1936–1939) led to a significant movement of people throughout the Catalan-speaking territories. However, the real second wave of immigration took place between 1950 and 1975, when hundreds of thousands of people moved into the Catalan-speaking areas attracted by the jobs available in the industrial and service sectors, especially in the growing area of tourism (Domingo 2014b). At this time, migration affected all Catalan-speaking societies, with the (relative) exception of la Franja and transformed the demographic landscape of the area dramatically (see table 1). By 1970, for instance, almost 4 out of 10 residents in Catalonia, and 2 out of 10 in the Valencian Country and the Balearic Islands had been born elsewhere in Spain, mostly in Castilian-speaking regions. A significant share of those born in Catalonia were children of Castilian-speaking immigrants.

Table 1: Percentage of residents in Catalonia, the Valencian Country and the Balearic Islands in 1970 according to birthplace (source: Domingo 2014b, 21: our elaboration).

	the same territory	rest of Spain	abroad	total
Catalonia	62.6	36.7	0.7	100
Valencian Country	77.0	22.2	0.8	100
Balearic Islands	79.2	19.6	1.2	100

Immigration in the 20th century introduced a qualitative sociolinguistic change. Until its arrival, for a majority of the population Castilian had been a distant vehicle of power strongly associated with public administration, the ruling classes, officialdom and cultured environments, yet not an everyday interpersonal experience, as proved by denominations such as *llengua del cadastre* ‘cadaster language’, or even *foraster* ‘foreigner’. However, as the 20th century progressed, with the arrival of thousands of newcomers to the Catalan-speaking societies, more and more locals started to encounter speakers of the hegemonic language in their everyday life, and therefore increasingly found themselves in the position of having to use the new *national* language. As a consequence, by the end of the century proficiency in the state languages was widespread among the whole local population.

The last two decades of the 20th century were relatively calm as far as migratory movements in the Catalan Countries are concerned, with the exception of the Balearic Islands, Andorra and Northern Catalonia, where immigration continued (Domingo 2014b). Language policies developed after the end of Franco’s dictatorship that sought to recover Catalan were basically addressed to this mixed Catalan/Castilian-language population (↗19 Language Law and Language Policies; ↗21 Catalan in the Mass Media: The Rise of Stylebooks).

But with the arrival of the 21st century, the real estate bubble attracted immigration from many different countries to virtually all Catalan-speaking territories: specifically, between 1998 and 2011, a total of 2,742,395 people moved to Catalonia (1,483,083 people), the Valencian Country (1,033,412 people) and the Balearic Islands (225,900 people) (Domingo 2014b, 23). These immigrants differed from those that arrived in the 20th century in at least two ways: on the one hand, this was *external* immigration, i.e. newcomers did not have Spanish citizenship. On the other hand, their origins were very heterogeneous in terms of nationality and linguistic repertoire, and most did not speak a language understood by locals. In Catalonia, only one third of immigrants came from Latin America, whereas in the Valencian Country and the Balearic Islands, the percentage of Latin Americans was lower (Domingo 2014b, 25).

The global financial crisis of 2008 badly hit the real estate industry and provoked a recession and severe unemployment, which in Spain exceeded 25%. As a consequence, immigration rates diminished dramatically after 2010 and emigration – very often led by former immigrants – rose significantly during the decade, especially in

some areas of the Valencian Country and Andorra (Domingo 2014b). Nevertheless, the majority of those who had arrived during the boom period remained. All of these population movements are reflected in the current demolinguistic setup of the Catalan-speaking territories.

4 The current demolinguistic situation of Catalan

4.1 A medium-sized European language

In global terms, and taking into account both the speakers living in the Catalan Countries and those living abroad, the total number of native Catalan speakers can be calculated at around 5.6 million (see table 2). The number of people claiming to speak and read it can be calculated as somewhere between 10–11 million. The figure increases to 12–13 million for those who can at least understand the language and goes down to 7–8 million as far as writing abilities are concerned.

Table 2: Number of native speakers of Catalan (alone or alongside other languages) and people who can use Catalan according to their place of residence, 2013–2015 (source: our elaboration, on the basis of Sorolla 2017a, EUL 2013–2015 and official censuses).

Place of residence	native speakers	can understand	can speak	can read	can write
Catalan-speaking territories	5,079,000	11,907,000	9,671,000	9,996,000	7,387,000
Spain (outside Cat.-sp. territories)	269,000	478,000	442,000	382,000	347,000
Other countries	225,000	398,000	371,000	368,000	203,000
Total	5,573,000	12,783,000	10,484,000	10,746,000	7,937,000

A peculiarity of Catalan demolinguistics is that the ratio between total speakers and native speakers is almost 2:1. In other words, there is a non-native speaker – a “new speaker” – for virtually every native speaker. This quite exceptional ratio is due to the combined effect of massive immigration and revitalization efforts, especially within primary and secondary education. Although not all new speakers use Catalan actively in the course of their everyday lives, a considerable percentage do. This massive presence of new speakers is obviously having an impact at all levels; not only does it enlarge the potential market for products in Catalan, but it also entails consequences for language structures, language practices and speakers’ linguistic beliefs.

4.2 Seven territories with their own demolinguistic dynamics

The fact that the traditional Catalan-speaking area is historically fragmented into a number of territories that do not form a single political or administrative unit has been crucial in the demolinguistic evolution of the language. Table 3 shows the percentage of people claiming to be proficient in Catalan in the most recent surveys in each one of these territories.²

Table 3: Percentage of people claiming to be proficient in Catalan in each territory (source: our elaboration from EUL 2013–2015).

	understand	can speak	can read	can write
Catalonia	94.3	80.4	82.4	60.4
Valencian Country (Catalan-speaking area)	77.7	56.4	57.2	38.3
Balearic Islands	96.8	80.5	83.5	61.9
Andorra	96.4	86.6	92.5	76.6
la Franja	94.1	80.2	74.7	41.2
l'Alguer	88.2	50.5	35.6	8.1
Northern Catalonia	61.1	35.4	39.2	14.3

According to this data, in four territories more than 80 % of the population claim to be able to speak Catalan: Andorra, the Balearic Islands, Catalonia and la Franja. Meanwhile, two territories have considerably lower figures: l'Alguer, where approximately half of the population claims to be able to speak the language, and Northern Catalonia, where just over one third (35 %) of the inhabitants claim to be able to speak Catalan. The figure from the Catalan-speaking area of the Valencian Country falls between these two groups, although the figure grows to 83.3 % if we include those who said they could speak the language *a little*.

The same classification holds for the ability to *understand* Catalan, although the percentages are much higher and, in fact, approach universal comprehension in four territories with results beyond 90 %, followed by l'Alguer, with 88 %. Comprehension is lower in Northern Catalonia (61 %). The fact that Catalan is a Romance language and therefore typologically similar to Castilian, French and Italian is doubtless a

² It is always risky to compare data coming from different surveys. In this case, data from all territories except Andorra and the Valencian Country come from the same dichotomic question where people were required to answer YES/NO to the question “Can you speak Catalan?” Results for Andorra include those self-evaluated themselves with “5” in a scale from 0 to 10. The Valencian results show those claiming to know Catalan *perfectly* and *rather well*, but not those claiming to know it *a little*.

factor that enhances language comprehension. The Valencian results oscillate between 77.7 % and 96.4 %.

Percentages for reading are very similar to those of speaking in the leading territories, although significant divergences are visible in Andorra, where more people can read than speak, and in la Franja, where the reverse is true. The results in l'Alguer and Northern Catalonia for language reading are again lower than in the rest of the language area and relatively similar. In the Valencian Country, the addition of *a little* places the figure at 83.8 %.

Finally, the ability to write Catalan is by far the least widespread. It should not be forgotten that, with the exception of Andorra – which unsurprisingly obtains the best results –, Catalan was hardly taught in schools until the late 1970s. As a result, most Catalan-speakers over 50 never had the opportunity to learn how to write their language. Today, the intensity of Catalan language teaching ranges from being a basic, compulsory subject in most territories, to being a mere option after regular schooling in others (↗20 Teaching and Learning of Catalan). As in the case of reading, la Franja ranks significantly below the three to four leading territories, bearing testimony to the fact that Catalan is neither official nor compulsory at school in that territory. The percentages in Northern Catalonia are much lower, even more so in l'Alguer. The figures in the Valencian Country reach 63.8 % when the three categories are included.

4.3 A complex, almost fractal language community

The description of the seven territories should not lead to the belief that each Catalan-speaking territory is linguistically homogeneous. On the contrary, internal differences are easy to observe. In Catalonia, the Valencian Country and the Balearic Islands, for instance, urban and touristic areas tend to exhibit lower percentages of proficiency – due to immigration and language shift – compared to inland areas and/or medium-sized urban centers and non-touristic areas, where Catalan is the predominant language of everyday life. Compare for instance the ability to speak Catalan (according to EUL 2013–2015) in the Alacant region (44 %) or in the city of Valencia (45.8 %) with results in the neighboring Alcoi-Gandia region (77.8 %) and in the region of Valencia city (77.2 %). Results are also quite different between Eivissa and Formentera (70.6 %) and Minorca (80.2 %), where tourism is less pervasive.

It must be said that demolinguistic differences do not stop at the regional level. Due to their different history in language shift and attraction of immigration during the last century, significant differences can be detected between neighboring *comarques*, neighboring cities, and even districts and neighborhoods. To give but one example, according to the 2011 census, 72 % of the population of Barcelona claimed to be able to speak Catalan, a percentage that was quite different from those obtained in the neighboring cities of Santa Coloma de Gramenet (51 %), Badalona (62 %) or Sant

Cugat del Vallès (82%). Differences between districts and neighbourhoods within one single city may also be very significant. To some extent, the current sociolinguistic structure of the Catalan-speaking territories may be described as “fractal” in the sense that it has “an irregular or fragmented shape at all scales of measurement between a greatest and smallest scale”.³

The proportion of Catalan speakers is strongly related to the proportion of native speakers and the proportion of people who identify Catalan as “their language”.⁴ Figure 1 shows these percentages for each territory, and allows for the distinction between two groups: those regions in which more people identify with Catalan than have it as their first language (as in Andorra, Catalonia and, to a lesser degree, la Balearic Islands); and those where both percentages are broadly similar (i.e. la Franja, l’Alguer and Northern Catalonia). In the former, the position of Catalan is strong enough to attract native speakers of other languages to identify with it, whereas in the latter group this is not the case.⁵

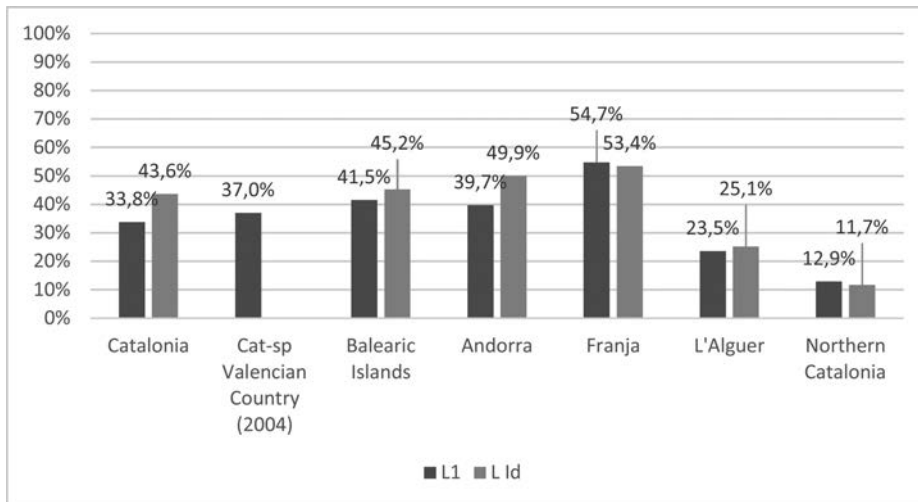


Figure 1: Percentage of people with Catalan (alone or alongside other languages) as their first language and as their language of identification, in each territory (sources: for most territories, EUL 2013–2015, cited as analyzed from by Vila/Sorolla 2017a; for the Valencian Country, Acadèmia Valenciana de la Llengua 2005, 90).

³ <http://www.dictionary.com/browse/fractal?s=t>

⁴ Demolinguistic research in the Catalan Countries has repeatedly proved that these two magnitudes, known as *first/initial language* and *language of identification* respectively, are correlated but still exhibit different behavior.

⁵ Unfortunately, language of identification is not available for the Valencian Country, and figures for first language are one decade older than L1 data from the other territories. Notice that the latter were collected in the midst of the last immigration wave. Current percentages are probably lower than those shown in the graph.

The demolinguistic configuration of each territory depends on the specific combination of several social variables, which adopt different characteristics in each territory and are very difficult to synthesize here (Querol et al. 2007; Vila et al. 2017; Cases-noves Ferrer 2012). Some of their most basic elements include:

- Family origin is by far the strongest determinant variable in predicting language proficiency and language use. Territories with higher levels of immigration, for instance Eivissa, Northern Catalonia or the industrial belts around the major cities, tend to obtain lower results for Catalan than those with less immigration such as la Franja, Minorca or the region of Alcoi-Gandia. In all territories, natives and their descendants are much closer to Catalan – in terms of language proficiency, use and identification – than immigrants and their descendants.
- Historical resistance against language shift is also a determinant variable in understanding language distribution. Despite immigration, Andorra and Catalonia usually exhibit better trends than the other territories, whereas l'Alguer and Northern Catalonia come in last.
- Socioeconomic and educational position tends to be a powerful explanatory variable for language distribution, but it is strongly connected with the local history of language shift. In territories where language shift has not taken place such as Andorra and Catalonia, Catalan tends to be better known and more widely used among people that are better educated, whereas lesser educated groups, to a large extent derived from migratory processes, tend to be less familiar with the language. On the contrary, in l'Alguer and Northern Catalonia, where language shift has taken place, Catalan tends to be mastered and used more by lesser educated people. The correlation of class and language is not so strong in the Valencian Country, where immigration has affected all social classes more evenly.
- In demolinguistic terms, sex and gender have very limited explanatory power as far as language proficiency and language use are concerned.
- Age is, next to origin, one of the most significant variables in predicting language proficiency and language behavior (see 4.5).

4.4 A complex pattern of multilingualism in language behavior

The previous sections have shown that Catalan is distributed in a very complex way across the different Catalan-speaking societies in terms of language proficiency, use as a first language and linguistic identification. Inevitably, this complex distribution results in an intricate pattern of language uses. Figure 2 displays language use in all territories in four social domains: (a) at current home, the most private domain for most people; (b) at large commercial establishments; (c) at banks and savings banks; and (d) with local authorities.

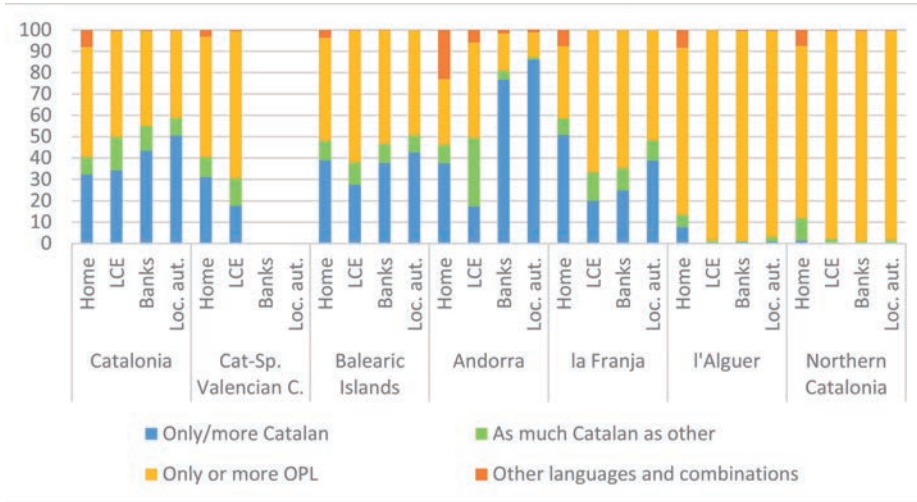


Figure 2: Percentage of people who use each language in four different situations in every Catalan-speaking territory (source: EUL 2013–2015, cited from Vila/Sorolla 2017 c).

OPL = “Other principal language”, i.e. Castilian (in Catalonia, the Valencian Country, the Balearic Islands, Andorra and la Franja), Italian in l’Alguer and French in Northern Catalonia.

At first glance, the graph demonstrates that the use of Catalan is significantly lower than language proficiency in all territories. The gap between proficiency and use is to be expected, since language skills are additive (one person may be very proficient in many languages at the same time), whereas language practices are in many respects a zero-sum variable. In this regard, language use is closer to language identification than to language proficiency.

A second look at the data in Figure 2 makes it clear that, in broad terms, language practices reproduce – in fact, intensify – the geographic distribution we saw in Section 4.1: Catalan is much more widely used in Catalonia, the Valencian Country, the Balearic Islands, Andorra and la Franja, than in l’Alguer and Northern Catalonia. In fact, the graph makes the degree of language minoritization of Catalan in the latter territories evident, for it is seldom used beyond private domains such as the homes of close friends (not shown in the graph).

A closer examination of the data provides new insights into the sociolinguistic reality of the Catalan-speaking territories. At home, Catalan is the predominant language only in la Franja and Andorra, whereas it is the second in the other territories. However, the position of Catalan is very diverse according to the territory and domain. In Catalonia and especially in Andorra, Catalan is more widely used outside the home than inside, i.e. it serves the role of a public language for people who do not use it at home – be it because it is not their first language or because, even if it is their L1, they speak another language at home probably because they are

married to monolingual speakers of other languages. In the Balearic Islands, the use of Catalan remains quite stable, signaling that, as far as the big picture is concerned, use of the language has so far not receded among native speakers but has not made progress among newcomers either. In the Valencian Country and la Franja, at-home speakers of Catalan cannot use their language in many public spheres and are therefore compelled to use Castilian instead.

Explaining each particular result requires a considerable amount of sociolinguistic information. To understand language practices in large establishments it is crucial to know whether a large percentage of a company's staff is made up of foreign migrants that most of the time are not competent in Catalan. The data from la Franja, for instance, must be understood taking into account that quite often their banks and local administration are manned by monolingual Castilian speakers from elsewhere in Aragon. In any case, the data confirm that today *diglossia* is not a crucial concept for describing the current state of language practices in the Catalan language territories, because none of the territories shows a neat and stable distribution of two languages between so-called "high" and "low" functions. Rather, the behavior detected in most territories corresponds to the dichotomy *natives vs non-natives* (and their descendants).

4.5 Intergenerational evolution

The comparison of language practices among successive generations tends to indicate temporal evolution. Intergenerational changes must of course be analyzed in connection with other phenomena such as migration, since differences may be due to population movements rather than to endogenous transformations. In general terms, languages which are more widely spoken by older generations than younger ones tend to be endangered, whereas those which gain speakers among the youngest cohorts tend to expand.

All Catalan-speaking territories share a basic component in their intergenerational profile: in all territories the figures diminish progressively from the oldest generation to, at least, the young adults' generation. This slope bears witness, in general terms, to the fact that the presence of speakers of other languages in this area is directly connected to two interrelated phenomena that took place and grew in intensity since the mid-20th century in particular: immigration and the nativization of the respective state languages, on the one hand, and language shift among locals towards these languages on the other. Yet despite general similarities, each territory displays a different profile according to its unique history (see Table 4)).

Table 4: Percentage of people claiming Catalan (alone or alongside other language/s) as their first language (L1) and as their language of identification (L Id) in the seven territories (source: our elaboration from EUL 2013–2015).

		15 to 29 years	30 to 44 years	45 to 64 years	65 or more	Total population
Catalonia	L1	40.7	29.2	32.2	36.4	33.8
	L Id	47.7	39.5	44.0	45.3	43.6
Valencian Country: <i>no comparable data are available</i>						
Balearic Islands	L1	42.8	33.3	41.4	58.2	41.5
	L Id	44.8	37.9	45.9	60.5	45.2
Andorra	L1	51.8	32.4	33.5	53.6	39.7
	L Id	51.4	43.1	46.4	69.4	49.9
la Franja	L1	37.7	45.8	61.8	69.0	54.7
	L Id	38.7	41.1	60.1	70.9	53.4
l'Alguer	L1	5.4	13.4	26.2	37.0	23.5
	L Id	20.4	11.7	27.1	36.4	25.1
Northern Catalonia	L1	3.9	4.1	14.3	24.5	12.9
	L Id	3.5	6.1	12.2	21.3	11.7

In Catalonia, the biggest and most highly populated territory, the intergenerational evolution is moderately positive for Catalan for both indexes. In the recent past, the combination of adverse language policies and numerous and repeated waves of immigration caused a dramatic reduction in the percentage of Catalan speakers which was clearly visible in 2013 (see Table 4). The percentage of Catalan speakers reached its lowest point among the generation born between 1969 and 1983, i.e. the baby boomers born to parents who had grown up during the second half of Franco's dictatorship, many of them arriving in the Catalan Countries from Southern Spain. The percentage of Catalan speakers in this generation was further reduced by the arrival of thousands of (young adult) immigrants during the first decade of the 21st century. Simultaneously, though, language policies in favor of Catalan in schools, public administration and mass media visibly increased the value of the language, not only for native speakers but also in the eyes of the children and grandchildren of former immigrants. The consequence of this was that a potential language shift among Catalan-speakers was averted, whereas a non-negligible proportion of non-Catalan L1 speakers adopted this language in their everyday life and even transmitted it to their children (Direcció General de Política Lingüística 2015). As a result, indexes

for Catalan in the youngest generation are not only significantly higher than that of the previous cohort; they are also higher than the mean, and in fact the highest result for all generations currently alive. A second factor is the relationship between having Catalan as a first language and a language of identification. In Catalonia, the results for language of identification (which is based on the personal decision of each individual regarding his or her personal affiliation with a language) are higher than those for Catalan as a first language, which was ultimately decided by their parents. In other words, the movement towards adopting Catalan is significant.

The data from the Balearic Islands are similar to those from Catalonia in that they slope downwards from the oldest group to young adults, but there is some growth amongst young people. The decrease is more abrupt than in Catalonia, reflecting two factors: firstly, immigration started later in the Balearic Islands than in Catalonia, which means that non-Catalan speakers are less common among the older generation, but very numerous among adults and young adults. Nevertheless, although growth among the younger generation suggests little language shift among locals, the small gap between first language and language of identification in all cohorts implies a limited capacity to attract non-native speakers to adopt Catalan. This is in stark contrast to Catalonia, where, despite percentages amongst youngsters still being far below those of the oldest generation, new speakers are adopting Catalan.

Andorra's profile is an exception to the general pattern of gradual generational decline in two senses. The percentage of Catalan falls dramatically between the oldest generation and mature adults, but then remains stable between this group and young adults, and bounces back again between the latter and the younger generation. Rather than a slope, the trend looks more like a basin. This profile seems strongly determined by the social and labor structure of Andorran society, which attracts significant numbers of temporary workers – hence the lower percentages of adults – but does not retain many of them.

The other three territories show a similar slope profile which is clearly indicative of a process of language shift in progress, although at different stages.

The results among the oldest generation in la Franja are the highest of all territories, bearing testimony to the fact that, until recently, this used to be a relatively rural and peripheral area more prone to emigration than to immigration, and entirely unaffected by language shift. But over recent decades, this area, where there is virtually no official recognition of Catalan, has entered into the late capitalist society, has become the place of residence of adults and young people from the rest of the autonomous community, and has attracted foreign immigrants that have learnt the official language rather than Catalan. As a result, the percentage of Catalan speakers has halved between the oldest and the youngest generations. Percentages of Catalan speakers in this generation now rank second to last out of all territories.

The data for l'Alguer are considerably worse for Catalan. In terms of first language, it still scores a significant 37 % among the oldest generation, but its slope is so pronounced that it reaches 5.4 % among the youngest generation. Interestingly, the

evolution of linguistic identification is much smoother, and shows a significant recovery between young adults and the youngest cohort that suggests that emotional links with the language can still be mobilized in significant sectors of Alguerese society.

The situation is even more delicate in Northern Catalonia. The point of departure among the oldest generation is already much lower than in the other territories – Northern Catalonia is a favored area for people from northern France to enjoy their retirement – and, although the speed of retreat seems to have reduced among the youngest generation, the percentage of both L1 and identification with Catalan are lower than in any other territory.

Besides immigration and language policies, two other elements seem crucial to understanding these differences in intergenerational evolution. On the one hand, intergenerational transmission of Catalan is significantly different in each territory; in Andorra and, to a lesser extent, in Catalonia, significant percentages of L1 speakers of other languages (basically Castilian) tend to adopt Catalan and transmit it to their children, whereas in the Balearic Islands the trend is less positive and in l'Alguer and Northern Catalonia it was native speakers of Catalan who interrupted family transmission between the first and third generations (Torres i Pla 2017). In fact, the breakdown of intergenerational transmission seems to have started recently in la Franja (Reyes et al. 2017; Sorolla 2017b). On the other hand, attitudes towards Catalan are very different in each territory. Using Catalan is much more common among young non-Catalan L1 Andorrans and Catalonians than in the other territories (Bretxa 2017).

5 Synthesis and conclusion

Fragmented into several territories under different political and administrative powers, Catalan has undergone a secular process of legal, political and cultural minoritization which has been combined, during the last century, with massive immigration all over its domain and, in some regions, language shift. As a consequence, in some of the regions where it has traditionally been spoken, Catalan is now severely threatened. Nevertheless, in spite of these difficulties, in global terms Catalan has survived until today as a vibrant medium-sized language community. Though always in a context of societal bilingualism or even multilingualism, the language is widely used in most of its territories and in most domains, enjoys relatively healthy intergenerational transmission, and, especially in some territories, namely Catalonia and Andorra, has a remarkable capacity to attract new speakers.

Each territory has followed a different historical path and shows a different demolinguistic profile. Catalonia, the largest and most populated territory, was the forerunner in attracting non-Catalan immigration in the first half of the 20th century and has continued to attract it in subsequent migratory waves. Catalanian society, though, has managed to counteract the ideology of Castilian as the national language

and avoided significant language shift towards the state language. In addition, it has managed to develop some active policies to secure the status of Catalan. As a consequence, although the position of Catalan in Catalonia has been shaken by the intensity of migratory movements, the language displays a considerable vitality in demolinguistic terms. Together with Andorra, whose independent status facilitates the adoption of policies that protect the language, Catalonia scores the highest in terms of the linguistic integration of newcomers among the Catalan-speaking territories.

The situation of Catalan is more complex in the Balearic Islands and the Valencian Country. On the one hand, in contrast with Catalonia and Andorra, a massive language shift within the family did take place in some areas, especially in Valencian urban centers. On the other hand, although immigration only started in the 1950s, the linguistic integration of newcomers has been much less widespread than in Catalonia or Andorra. As a result, the use of Catalan is more restricted to native speakers and less prominent in formal domains, where Catalan speakers often feel compelled to adapt linguistically to their monolingual interlocutors.

The demolinguistic situation of la Franja is contradictory and undergoing rapid change. A rural area largely inhabited by an aged population which was sheltered from immigration and language shift until the end of the 20th century, it is still one of the demographic strongholds of native Catalan speakers. But the lack of official status is proving lethal for the language, which shows signs of receding and succumbing to language shift.

Finally, the situation of Catalan in l'Alguer and Northern Catalonia is much more precarious. The combination of Catalan's lack of official recognition, the peripheral socioeconomic position of its speakers, the establishment of state language speakers and the interruption of intergenerational transmission in the 20th century has led the language to a position of extreme minoritization, especially among the youngest generations.

6 Bibliography

- Acadèmia Valenciana de la Llengua (2005), *Llibre blanc de l'ús del valencià – 1. Enquesta sobre la situació social del valencià*. 2004, València, Acadèmia Valenciana de la Llengua.
- Anguera, Pere (1997), *El català al segle XIX. De llengua del poble a llengua nacional*, Barcelona, Empúries.
- Bernardó, Domènec/Rieu, Bernat (1976), *Diglossia a Catalunya Nord*, Treballs de Sociolingüística Catalana 1, 55–62.
- Bourdieu, Pierre (1982), *Ce que parler veut dire. L'économie des échanges linguistiques*, Paris, Fayard.
- Bretxa, Vanessa (2017), *Actituds i representacions lingüístiques*, in: F. Xavier Vila et al., *Llengua i societat als territoris de llengua catalana 2013–2015*, Barcelona, Generalitat de Catalunya, Departament de Cultura, Direcció General de Política Lingüística.

- Cabré Pla, Anna (1999), *El sistema català de reproducció. 100 anys de singularitat demogràfica*, Barcelona, Enciclopèdia catalana.
- Casesnoves Ferrer, Raquel (2012), *El valencià en 25 anys. Com són els seus parlants?*, Revista de Llengua i Dret/Journal of Language and Law 58, 111–136.
- Chessa, Enrico (2011), *Another case of language death? The intergenerational transmission of Catalan in Alghero*, Ph. D. thesis, London, Queen Mary, University of London, <https://qmro.qmul.ac.uk/jspui/bitstream/123456789/2502/1/CHESSAAnotherCase2012.pdf> (last accessed: 29.06.2018).
- de Gabriel, Narciso (1997), *Alfabetización y escolarización en España (1887–1950)*, Revista de Educación 314, 217–243, <http://www.mecd.gob.es/dctm/revista-de-educacion/articulosre314/re3141100462.pdf?documentId=0901e72b81272c6b> (last accessed: 29.06.2018).
- Direcció General de Política Lingüística (2015), *L'Enquesta d'usos lingüístics de la població 2013. Resum dels factors clau*, Barcelona, Generalitat de Catalunya, Departament de Cultura, Direcció General de Política Lingüística, http://llengua.gencat.cat/web/.content/documents/publicacions/btpl/arxius/15_EULP2013_factors_clau.pdf (last accessed: 29/06/2018).
- Direcció General de Política Lingüística (2016), *Coneixements, usos i representacions del català al conjunt del domini lingüístic*, Barcelona, Generalitat de Catalunya, Departament de Cultura, Direcció General de Política Lingüística, http://llengua.gencat.cat/web/.content/documents/dadesestudis/altres/arxius/Coneixements_usos_representacions_catala_conjunt_domini_linguistic.pdf (last accessed: 29.06.2018).
- Direcció General de Política Lingüística (ed.) (2018) *Anàlisi de l'Enquesta d'usos lingüístics de la població a Catalunya 2013. Volum 1: Coneixements, usos, transmissió i actituds lingüístics*. Barcelona, Generalitat de Catalunya, Departament de Cultura, Direcció General de Política Lingüística, http://llengua.gencat.cat/web/.content/documents/publicacions/btpl/arxius/23_EULP_Catalunya_2013-vol.1.pdf (last accessed: 29.06.2018).
- Domingo, Andreu (2014a), *Catalunya al mirall de la immigració*, Barcelona, L'Avenç.
- Domingo, Andreu (2014b), *Balanç i perspectiva demogràfica del conjunt dels territoris de llengua catalana*, in: Direcció General de Política Lingüística, *Jornada sobre Llengua i Societat als Territoris de Parla Catalana*, Barcelona, Generalitat de Catalunya, Departament de Cultura, 7–69.
- Escribano, Daniel (2015), *El règim d'oficialitat lingüística a Catalunya durant la Segona República espanyola*, Recerques. Història. Economia. Cultura 70, 125–159.
- Escudero, Jean-Paul (2000), *Algunes dades sobre els gitanos de Catalunya i la seva antiga llengua*, Treballs de Sociolingüística Catalana 14/15, 53–58.
- Espinet, Francesc (2001), *Algunes consideracions sobre classes socials i ideologia lingüística a Catalunya*, in: Toni Mollà, *Ideologia i conflicte lingüístic*, Alzira, Bromera, 91–144.
- Fabà, Albert/Montoya, Brauli (2012), *La transmissió lingüística intergeneracional del català al País Valencià. Una perspectiva territorial*, Caplletra. Revista Internacional de Filologia 53, 211–231, <https://www.raco.cat/index.php/Caplletra/article/view/267984/363357> (last accessed: 29.06.2018).
- Heller, Monica/Duchêne, Alexandre (2012), *Pride and Profit. Changing Discourses of Language, Capital and Nation-State*, in: Alexandre Duchêne/Monica Heller (edd.), *Language in Late Capitalism: Pride and Profit*, New York, Routledge, 3–21.
- Marfany, Joan-Lluís (2001), *La llengua maltractada: el castellà i el català a Catalunya del segle XVI al segle XIX*, Barcelona, Empúries.
- McRoberts, Kenneth (2001), *Catalonia: Nation Building without a State*, Oxford, Oxford University Press.
- Montoya Abad, Brauli (1986), *Variació i desplaçament de llengües a Elda i a Oriola durant l'edat moderna*, Alacant, Institut d'Estudis Juan Gil-Albert.
- Montoya Abad, Brauli (1996), *Alacant: la llengua interrompuda*, València, Denes 10.
- Montoya Abad, Brauli (2009), *La doble frontera entre les llengües (social i territorial): el Regne de València (Espanya) en els segles XV i XVI*, in: Teddy Arnavielle/Christian Camps (edd.), *Discours*

- et savoirs sur les langues, ancien(ne)s et modernes, dans l'aire méditerranéenne*, Paris, L'Harmattan, 167–180.
- Montoya Abad, Brauli/Mas i Miralles, Antoni (2011), *La transmissió familiar del valencià*, València, Acadèmia Valenciana de la Llengua.
- Nadal, Joaquim/Giralt, Eugeni (1960), *La population catalane de 1553 à 1717: l'immigration française et les autres facteurs de son développement*, Paris, SEVPEN.
- Ninyoles, Rafael Lluís (1969), *Conflicte lingüístic valencià*, València, Tres i Quatre.
- Pueyo, Miquel (1996), *Tres escoles per als catalans: minorització lingüística i implantació escolar a Itàlia, França i Espanya*, Lleida, Pagès.
- Querol, Ernest, et al. (2007), *Llengua i societat als territoris de parla catalana a l'inici del segle XXI. L'Alguer, Andorra, Catalunya, Catalunya Nord, la Franja, Illes Balears i Comunitat Valenciana*, Barcelona, Secretaria de Política Lingüística, Departament de Cultura, Generalitat de Catalunya.
- Reyes, Anchel, et al. (2017), *L'aragonés y lo catalán en l'actualitat*, Zaragoza, Universidad de Zaragoza, <https://zaguan.unizar.es/record/60448/files/BOOK-2017-009.pdf> (last accessed: 29.06.2018).
- Sorolla, Natxo (2017a), *Context demogràfic i econòmic. L'evolució de la comunitat lingüística*, in: Xarxa Cruscat, *IX Informe sobre la situació de la llengua catalana (2016)*, Barcelona, Observatori de la Llengua, <http://blogs.iec.cat/cruscat/publicacions/informe/> (last accessed: 29.06.2018).
- Sorolla, Natxo (2017b), *La llengua a la llar i la transmissió lingüística intergeneracional*, in: Direcció General de Política Lingüística, *Els usos lingüístics a la Franja, 2014*, Barcelona, Generalitat de Catalunya, Departament de Cultura, http://llengua.gencat.cat/web/.content/documents/publicacions/btp/arxiu/19_Usos_linguistics_Franja_2014.pdf (last accessed: 18.03.2019).
- Torres i Pla, Joaquim (2017), *Anàlisi de les Enquestes d'usos lingüístics als territoris de llengua catalana 2013–2015. Factors clau*, in: F. Xavier Vila et al., *Llengua i societat als territoris de llengua catalana 2013–2015*, Barcelona, Generalitat de Catalunya, Departament de Cultura, Direcció General de Política Lingüística.
- Vila, F. Xavier (2014), *Language Policy, Management and Planning*, in: Christiane Fäcke (ed.), *Manual of Language Acquisition*, Berlin/Boston, De Gruyter, 50–68.
- Vila, F. Xavier/Sorolla, Natxo (2017a), *Llengua inicial, llengua d'identificació i llengua inicial als territoris de llengua catalana*, in: F. Xavier Vila et al., *Anàlisi de les Enquestes d'usos lingüístics als territoris de llengua catalana 2013–2015. Factors clau*, Barcelona, Generalitat de Catalunya, Departament de Cultura, Direcció General de Política Lingüística.
- Vila, F. Xavier/Sorolla, Natxo (2017b), *Els usos lingüístics interpersonals*, in: F. Xavier Vila et al., *Llengua i societat als territoris de llengua catalana 2013–2015*, Barcelona, Generalitat de Catalunya, Departament de Cultura, Direcció General de Política Lingüística.
- Vila, F. Xavier/Sorolla, Natxo (2017c), *Els usos lingüístics segons els interlocutors i els àmbits*, in: F. Xavier Vila et al., *Llengua i societat als territoris de llengua catalana 2013–2015*, Barcelona, Generalitat de Catalunya, Departament de Cultura, Direcció General de Política Lingüística.

Eva Pons

19 Language Law and Language Policies

Abstract: In this chapter we analyze the main features of language laws and policies related to the status and usage of the Catalan language. The first part introduces the main characteristics of the different general legal frameworks: Catalan as official language in Andorra and the Autonomous Communities in Spain; Catalan as a regional language in France; and Catalan as a linguistic minority in Italy. In the second part, the analysis focuses on the content of the legal protections and language policies developed in various traditionally Catalan-speaking countries and regions, that is, Andorra, Aragon, l'Alguer, El Carxe (in Murcia), the Valencian Community, the Balearic Islands and Catalonia. Finally, we reach some general conclusions regarding the impact and relevance of formal legal status on language vitality, taking into account recent developments in legislation and jurisprudence that have restricted language policies aimed at protecting and promoting the Catalan language in Catalonia and other Autonomous Communities in Spain where it enjoys joint official status with Castilian.

Keywords: official language, linguistic minority, regional language, language rights, Catalan language

1 Introduction

The Catalan language represents the ninth biggest language in terms of speakers within the European Union. Catalan can be considered a medium-sized language, with corresponding implications in terms of the design and the scope of language policies applied to it (Milian 2012; Bastardas/Boix/Torrens 2018). Nevertheless, in legal and political-administrative terms, Catalan is not included in the 24 official languages of the European Union. In fact, with the exception of Andorra, a general feature of language policies concerning Catalan is its legal coexistence with another official state language. Despite differences between current linguistic models in the Catalan-speaking areas in France, Italy and Spain, in all of these countries legal protections are imposed at the regional level and generally do not reach the state level.

Related to the situation just described, a relevant feature of language law and policies concerning the Catalan language is the level of development and sophistication achieved (Pons 2011). Especially in those territories where Catalan has official status (in the Autonomous Communities of Spain), constitutional and legal provisions, interpreted by a large jurisprudence, serve to regulate language rights and

<https://doi.org/10.1515/9783110450408-027>

language usage in many public and social sectors. Public authorities have also developed a considerable number of language policies to promote the use of the Catalan language and encourage its spread among the population, with the ultimate aim of *linguistic normalization*. The process of linguistic normalization ‘involves the adoption of a series of legislative, political and administrative measures, applicable in different moments in time aimed at promoting a specific language that is considered to suffer from discrimination and/or requires special protection, or which, in any case, finds itself in a disadvantageous or undesirable position’ (Vernet/Pons 2011, 59). This situation is very different from other territories where Catalan has historically been used up until the present day but where the language does not enjoy official status (in Northern Catalonia in France or in l’Alguer in Italy), but *regional* or minority status.

In this chapter, we analyse the main features of language law and policies related to the status and usage of the Catalan language. The first part presents the main characteristics of the different general legal frameworks. The analysis then focuses on the content of the legal protections and language policies in different countries and regions. Finally, some general conclusions are reached regarding the impact and relevance of formal legal status on language vitality.

2 One language, varying legal status

In the era of globalization, linguistic issues remain a typical concern of sovereignty for states as a crucial element of their constitutional identity (Pons 2015). Although an increasing number of international standards have been developed in this field since the 1990s, they tend to focus on the linguistic dimension of human rights and the protection of minorities as a special dimension of human rights (Pons 2017). Universal linguistic standards, as interpreted by international bodies, do not concretely lay out how linguistic pluralism is to be established at the state level (Patten/Kymlicka 2003) and, in most cases, exhibit a limited capacity to counteract the assimilationist policies of individual countries (Paz 2013).

The most advanced enforcing instruments can be found in the European regional context, including those created by the Council of Europe, such as the Framework Convention for the Protection of National Minorities (FCNM, 1995) or the European Charter for Regional or Minority Languages (ECRML, 1992), the latter taking a new approach to linguistic diversity as a good or a value to be protected (see primarily Nogueira/Ruiz/Urrutia 2012). However, in both cases, these instruments have to be ratified voluntarily by states and one could also question whether they offer a true counterbalance to a political order effectively controlled by national governments (Kraus 2018).

Catalan is protected by the ECRML in Spain, under the general objectives of Part II and also Part III in some territories (in the Autonomous Communities where this language has official status). In contrast, neither Italy nor France have ratified this

instrument, even after having signed it. The French Constitutional Council (FCC), declared the ECRML incompatible with the declaration of French as the language of the Republic in the constitution (FCC Ruling of 15 June 1999).

2.1 Catalan as an official language

Official status is the highest legal status that can be assigned to a language. But *official language* or *official recognition* are contingent notions, the meaning of which only becomes concrete when established within the legal framework and related language policies in a given context (Pons 2017).

Catalan is the only official language of Andorra, a microstate located in the Pyrenees. The language has been historically present in the public life of this territory, which was subject to different jurisdictions until 1814, when it was recognized as a state under a condominium of the bishop of Urgell (Catalonia, Spain) and France. Under Article 2 of the Andorran Constitution (AC) of 1993, the official status of Catalan is effective in all areas of law. Nevertheless, in sociolinguistic terms, the political (two coprinces, the president of the French Republic and the bishop of Urgell, act as heads of state of Andorra) and linguistic influence of two neighbouring countries, combined with high immigration rates, have restricted language vitality in everyday life.

In Spain, Article 3 of the Spanish Constitution of 1978 (SC), established during the democratic transition following Franco's death (1975), hints at a certain degree of official recognition of linguistic and cultural diversity. The Catalan language was granted official status in the Statutes of Autonomy of the Balearic Islands, Catalonia and the Valencian Community (1979–1983), where Castilian – in as much as it is the state language – is also official. This status means that communications are valid and effective in any language declared as official, without the need for translations or bilingual publications (CC Judgment 82/1986, of 22 June). Through their institutions, the Autonomous Communities can pass legislation in order to determine the effects of this official status and to *normalize* the autochthonous language in different social sectors. However, central institutions can also regulate linguistic rights and usage through sector-based norms in those areas subject to competence of the state (Vernet et al. 2003; Cabellos 2008).

In the Spanish Autonomic State, autonomous policies related to Catalan are generally subject to resistance from the central government due to political trends towards centralization and homogenization (Vernet 2007; Milian 2009). More recently, the Constitutional Court Judgment 31/2010 of 28 June on the new Statute of Autonomy of Catalonia of 2006 introduced new limits on the official status of Catalan and a devaluation of the idea of 'linguistic autonomy' (Jou 2011; Segura 2011; Pons 2013; Milian 2016). The political-administrative division of the historical Catalan territories in the different Autonomous Communities also involves denying this official status in some Catalan-speaking areas of Aragon and Murcia, where the language

remains legally protected as a part of the cultural richness of the country under Article 3.3 SC (Arzoz 2009).

2.2 Catalan as a linguistic minority

The concept of *linguistic minority* implies an inferior, non-official legal status for minority languages. Similar to official status, this notion must be shaped by the concrete legislation and policies associated with this minority status (Toso 2008). In Italy, the Constitution of 1947 stipulates that ‘the Republic safeguards linguistic minorities by means of appropriate measures’ (Article 6 IC). Initially, German and French were identified as ‘official minorities’ in specific regions. Only fifty years later a general legal system was established (Vernet et al. 2003).

The Law 482/1999 of 15 December regarding the protection of historical linguistic minorities declares, after confirming the official status of Italian, that the Republic values the linguistic and cultural heritage of the Italian language, whilst at the same time promoting the recognition of the languages and cultures protected by this law (Art. 1). The legal provisions regarding the use of protected languages negate any prerogatives of official status (e.g. the legal validity of communications and guaranteed linguistic rights). In addition, the implementation of regulatory measures in the traditional areas of historic minorities depends on special agreements adopted by local bodies (such as provincial councils, with the participation of municipal councils and the local population). Catalan is mentioned by Article 2 of Law 482/1999 as the language traditionally spoken in the municipality of l’Alguer in Sardinia (Caria 2014). At the regional level, the Sardinian Law 26/1997 of 15 October granted Catalan minority status under a minority regime primarily devoted to the protection of the Sardinian language. Recently, this Law has been abrogated by Law 22/2018 of 3 July on Discipline of Regional Linguistic Policy, which substantially reinforces the minority status of Catalan in l’Alguer as part of the ‘intangible heritage of the Region’, deserving of equal measures of protection, valorization and promotion as the Sardinian language, according to Article 2 (an inferior level of protection is granted to Sassarese, Gallurese and Tabarchino).

2.3 Catalan as regional language

The notion of *regional language* refers here to the internal consideration granted in France to autochthonous languages other than French. While there is political opposition to the recognition of a ‘minority’ status from the linguistic groups concerned, the legal implications of the notion of *regional language* remain ambiguous or uncertain (Zabaleta 2015). The constitutional amendment introduced in 2008, according to which ‘regional languages belong to the heritage of France’ (Article 75-I FC), seems to

contribute little to changing the situation. This new precept was the subject of criticism due to its view of languages as ‘goods to preserve in a museum’, and not as instruments of social communication (Lasagabaster 2017). Afterwards, this constitutional provision did not experience any legal or judicial development. But the previous reform of the Constitution in 1992, which establishes French as ‘the language of the Republic’ (Article 2), sets out a negative definition of regional languages, which are tolerated in private contexts and in some specific areas of the public sphere, but do not have access to the Republic, citizenship or the law (FCC Ruling of 15 June 1999).

The subjection of the Catalan language to this legal framework denies any guarantee for its use in public administration in the historical territory of Northern Catalonia, covered by the Oriental Pyrenees Department, which includes the traditionally Catalan-speaking districts of Roussillon, Vallespir, Conflent, Cerdagne and Capcir. The only way this legal situation can evolve in a positive and significant manner (Zabaleta 2015) is through another constitutional reform that allows the ratification of the ECRML, as proposed in 2014 by the French Parliament.

3 Andorra

Under the Andorran Constitution of 1993, the official status of Catalan was established by the law governing the use of the official language (hereafter referred to as the LOULO), adopted on 16 December 1999. The preamble to this law states that the language is one of the main factors defining the identity of Andorra. It also argues that a number of factors such as ‘the proximity of two languages of major demographic proportions, the tradition of teaching these two languages in the Principality of Andorra, the impact of the mass media and, more recently, the high levels of immigration linked to the Andorran economy and society in recent decades may endanger the vitality of our language’. Therefore, the public authorities in Andorra have continued to pass various linguistic norms to ensure that, as a state language and as a means of collective expression among Andorran society, ‘Catalan has to be the language of general use’ (Preamble of the LOULO). This law shares a *normalizing* aim with other linguistic legislation (see below 3.6 to 3.8), to the extent that it seeks to prevent a decline in the knowledge and use of a native language, referred to in the preamble as ‘our language’ (Vernet/Pons 2011).

The general goals, set up in Article 2 of the LOULO, include guaranteeing the official use of Catalan, encouraging its spread throughout Andorran society and preserving and guaranteeing its use in areas of public relevance, such as education, media and social and cultural activities (Art. 2 LOULO). General language-related rights are detailed under Article 3, which states that: ‘Everyone has the right to be addressed and replied to in Catalan in their oral and written relations with any public authority and the entities and bodies reporting to it; with the healthcare and social

services.’ Additionally, Article 4 indicates that all Andorrans have the duty to understand the Catalan language and use it in the cases foreseen by law (the Qualified Law on Nationality of 5 October 1995 establishes the knowledge of the Catalan language as a requirement for naturalization, Article 36).

Since the LOULO was adopted, the use of Catalan has been assured in various public sectors. Title I of the act (‘The Official Language’) establishes that ‘Catalan is the language employed by all public institutions of Andorra, in accordance with the law’, and is therefore the language of ‘laws, regulations and all other official texts; of all administrative and judicial activities and, in general, of the authorities, public companies and para-public entities’ (Art. 8). Other parts of the act cover more specific areas of the use of the official language: the toponymy and its use in signage, advertising, labelling or other documentation (Art. 14); its knowledge and use by staff at social or health facilities, with possible exceptions adopted by the government (Art. 18); public transport, while allowing the use of other languages when providing individualized services for users (Art. 19); and broadcasting and advertising in the media of Andorra, which may also have programmes or sections in other languages (Art. 25, pending the approval of the foreseen regulations). Title V of the LOULO contemplates a system of penalties in case of infringement of these legal provisions (Art. 36 to 41).

In practice, the Andorran authorities, at both the and local level – the *comuns*, public bodies governing the local districts known as *parròquies* –, use Catalan almost exclusively in their internal processes and in written communications with the public, and also in the majority of oral communications (Marí 2011), whilst also being able to turn to ‘the use of other languages with regard to foreigners who do not understand Catalan’ (Art. 8.1 LOULO). The official journal appears only in Catalan (Law 25/2014 of 30 October). Documents addressed to any Andorran authority must be submitted in the official language, but in cases where this poses reasonable difficulty ‘the relevant authority may dispense with this obligation, unless it is planned to make (the documents) public’ (Art. 8.1 LOULO). Special provisions are contemplated for international relations, which shall be governed by international regulations, granting an official Catalan version of international instruments to be signed and stamped by the Andorran authorities (Art. 8.2).

In economic and commercial sectors, the shortcomings are more evident. Hence the explicit legal provisions which aim at extending the use of Catalan to certain socio-economic domains (Art. 15, 16, 17, 21, 21 and 22 of the LOULO) including customer service (Art. 20). In the area of education, a particular compromise is visible in the coexistence of multilingual Andorran schools, with an increased use of Catalan (law of 9 June 1994 governing the Andorran educational system), with special legislation and international treaties governing how languages are taught in the French and Spanish schools, in which Catalan is only a teaching subject and not a language of instruction.

The Government of Andorra is legally obliged to oversee the use of the official language and, in general, to foster its use in all kinds of activities (Art. 31). The

Language Policy Service of the Andorran government organizes (officially certified) language courses, carries out linguistic inspections, generates statistics and other instruments of language planning and promotes the use of the Catalan language in specific social or cultural sectors. Concerning transnational exchanges and agreements to promote the language (Art. 35 LOULO), Andorra participates in the Ramon Llull Foundation, located in its territory, which is made up of the General Council of the Oriental Pyrenees Department, l'Alguer, the Network of Valencian Cities and the Ramon Llull Institute (a consortium created in 2002 by the Governments of Catalonia and the Balearic Islands in order to promote the Catalan language and culture abroad). Andorra's international cooperation is guided by a pragmatic approach, with the country holding membership of both the International Organization of the *Francophonie* and the Ibero-American Conference.

4 L'Alguer

According to the Italian Law 482/1999, the legal framework of Catalan in the municipality of l'Alguer includes measures in the following sectors: the non-compulsory use of the minority language as a medium of instruction and its teaching in pre-school, primary and secondary education, respecting general rules on education, and in accordance with the pedagogical and organizational methods adopted by the educational centres, granted the agreement of families, as well as linguistic training for adults (Arts. 4 and 5); specific language courses in universities (Art. 6); the use in the activity of municipal councils and the publication of official acts or resolutions, providing a translation for any other member who does not understand the minority language (Arts. 7 and 8, which restrict the official validity to the Italian version); the use in oral and written communication with the public authorities of the corresponding municipality, excluding the police and the armed forces (Art. 9); place names (Art. 10); recording names or surnames on the civil register (Art. 11), and the possibility of broadcasting in the public media in the protected language (Arts. 12 and 14).

Furthermore, Law 482/1999 empowered regional and provincial institutions to adopt more favourable provisions regarding these languages (Art. 13). In the case of Catalan, the Sardinian Law 26 of 15 October 1997 for the promotion and recognition of the Sardinian culture and language initially establishes that: 'The same value attributed to the Sardinian culture and language is recognized, with regard to the corresponding region, for the Catalan culture and language of l'Alguer' (Art. 2.4). By virtue of this equivalent status, the provisions of Title V of this act, referring to the use of the Sardinian language by the public authorities, i.e. in local collegiate bodies, in correspondence and oral communications between the public and the regional and local authorities (Art. 23.3) and in place names (Art. 24), are equally applicable to the Catalan of l'Alguer. These legal undertakings have translated into budgetary support from regional programmes for certain cultural initiatives concerning Catalan, educa-

tion and digital media. But specific recognition of Catalan within a legal text mainly aimed at promoting the Sardinian regional culture confers only a marginal character to this political support (Marí 2011).

The new Regional Law 22/2018 of 3 July, on language policy, seems to provide and answer to previous criticisms, by including an explicit regulation of Catalan in l'Alguer throughout the legal text, on an equal basis with Sardinian. These dispositions cover: political and administrative measures of protection, valorization, promotion (Art. 3–5); linguistic certification (Art. 9); language uses in Administration (Art. 10–12); the teaching of Catalan language and its use as vehicular language of all subjects of the curriculum in kindergarten, primary schools and secondary schools of first degree of l'Alguer, according to specific modalities corresponding to centre and degree (Art. 17); extracurricular laboratory in Catalan (Art. 19); linguistic proficiency for teaching staff, with a minimum of C1 (Art. 20); public interventions in the field of information, publishing and new technologies, including the agreement on a cultural and linguistic quota in radio and television concession (Arts. 22–23), and collaboration with universities and private sector (Arts. 24–26).

At a local level, the Statute of the Municipality of 1996 stated that 'The municipality [...] shall carry out the necessary actions to safeguard the language of l'Alguer' (Article 8). Adopted before a higher-ranking legal basis was established, this Statute highlights the cultural interest in the protection, promotion and dissemination of Catalan in its local variety, but makes no provision for the administrative use of the language. After its reform, the current Statute ensures support to 'all initiatives directed at the knowledge and use of the Catalan language, in its Alguerese variety', the use of bilingual Catalan-Italian place names, and the definition of the Municipality as 'institutional referent' for language policy, which subsidizes associations operating in that field (Article 9). In practice, the municipality offers information in Catalan on its website and has Catalan-speaking staff catering to tourists from other Catalan territories. The municipal action for language promotion, also with the cooperation of social organizations and the institutional delegation of the *Generalitat de Catalunya* (that has been recently suppressed, see 10 below), includes initiatives for Catalan courses for adults, cultural events or publications in Catalan (Observatori de la Llengua Catalana 2015). Article 4 of Regional Law 22/2018 establishes the obligation to adapt municipal statutes and regulations to Law 482/1999 and to its own provisions and recognizes the competences of local authorities and their collaboration with other administrations or public or private institutions, national and international (see also Art. 7, which fosters stable relations with other minority language communities), in order to spread the use of Catalan in private and public sectors. Social participation in defining linguistic measures is articulated through the *Consulta cívica per les polítiques lingüístiques del català de l'Alguer* (integrated by 11 associations, has met firstly in 18 June 2018), and has the technical support of a Catalan linguistic service (Art. 11).

5 Northern Catalonia

The exclusive official status of French and its compulsory use in almost all public sectors in France, as stated by the Law of 4 August 1994 (the *Loi Toubon*), reserves a marginal role for legislation on the use of Catalan as a regional language. The French State Council has identified education, media, research and translations as the areas where the use of other languages – including regional languages – is legally permitted, but still excludes any conceptualization of these uses as linguistic rights (Zabaleta 2015).

In education, the Act of 11 January 1951 (the *Loi Deixonne*, which was reformed in 1975, 1984 and 1989 and recently repealed) introduced the teaching of the regional language and the local dialects. The legislation in force establishes that the public education authorities have to provide the appropriate means to encourage the learning of local languages and dialects within the regions in which they are used (see Article L312–10 of the Education Code, introduced in 2013) and opens the door for the use of Catalan as a medium of instruction (in bilingual education programs) and not only as a subject of study. Under this legislation, optional Catalan courses are offered from pre-school to secondary education in certain state schools and colleges, with the compulsory teaching of the regional language being considered as contrary to the French constitution (FCC Decision 2011–130 QPC 20 May 2011, case *Madame Cécile Joref*). However, the breach of legislative mandates has prompted private initiatives to fulfil the educational needs in terms of the “regional language”. This was the origin of nest schools such as Bressola (‘cradle’) and Arrels (‘roots’) (Observatori de la Llengua 2015). Arrels dates back to 1995 and is still the only public school to offer education in Catalan as a medium of instruction. Some legal questions have arisen over the compatibility of immersion schools with French legislation (for example, students still have to pass the official exams in French), but lack of financial support is one of the most prominent difficulties such initiatives have faced (Lasagabaster 2017).

At the sub-state level, local and the regional institutions recognize the cultural Catalan identity of Northern Catalonia through public initiatives that foster the use of the language in some limited areas of public and social life. In 2007 the *Conseil Général des Pyrénées-Orientales* passed a ‘Catalan Language Charter’, in which this body ‘officially recognises Catalan as a language of the department alongside the French language’ (Art. 1) and accepts the Institut d’Estudis Catalans as a language academy. Accepting the declarative scope of the document as an expression of the political support to the local and regional initiatives concerning Catalan, its content refers to the use of the language in bilingual signs, in the services of the department – including communications with the public –, in the documents and publications produced (invitations, newsletters, programs, magazines, web pages, etc.) and in the public and private media. For example, the Council of Perpignan, the main city in the area, signed a similar charter in 2010 and has a website on which the Catalan language has a considerable presence (nevertheless, the section dedicated to the administrative

procedures only offers the possibility of using Catalan alongside French in a limited number of procedures).

6 El Carxe (Murcia)

The Spanish region of Murcia includes a Catalan-speaking area formed by the three municipalities of El Carxe (Iecla – Yecla in Castilian –, Jumella – Jumilla – and Favanella – Albanilla –), on the border with the Valencian Community. Although there is no specific reference to Catalan in the region's current legislation, the Statute of Autonomy of Murcia of 1982 states: 'The Autonomous Community [...] shall protect and promote the cultural peculiarities, as well as its popular customs and traditions, respecting the local and regional variants in all cases'. As a constitutional framework for this provision, Article 3.3 SC determines that 'the richness of the different language forms in Spain is a cultural heritage and shall be protected and respected in a special way'. The Catalan of Murcia, as a legally, though not officially protected language, is included under Part II of ECRML (Art. 7), which contains the objectives of protection and promotion of the language (Council of Europe 2012 and 2016).

In the absence of linguistic legislation, the regional Parliament has passed a number of laws that could be interpreted or developed as a source of protection and promotion of the Catalan language (Navarro 2018a; 2018b). For example, the legislation on cultural heritage, which considers the language as an intangible cultural heritage (Law 4/2007 of 16 March on the cultural heritage of the region of Murcia), or the legislation seeking to promote Murcian identity through the dissemination and awareness of its specific cultural and linguistic values in schools (Law 6/1998 of 30 November on the education councils), could be developed in a specific way in the Catalan-speaking area of El Carxe. In practice, general ignorance of this linguistic reality in the Region of Murcia as a whole has prevented until now the adoption of any political or administrative measures aimed at recognizing the specific language of El Carxe. The Catalan language has no public use in the local administration of the three municipalities, nor in the *pedanías* (minor local bodies), and only sparse manifestations of this particular linguistic identity are expressed in the cultural sphere. Possible examples of 'resolute action' – as required by Article 7.1.c ECRML – with respect to the linguistic needs of El Carxe, as a territory of Murcia with strong geographical and cultural connections with the neighbouring Valencian municipalities (Article 7.1.e ECRML), included financial support from the Valencian Community for Catalan courses and other linguistic initiatives (see the Valencian Government Order 65/2016 of 18 October, followed by others in 2017 and 2018).

7 La Franja de Ponent (Aragon)

The framework provided by the general protection derived from Article 3.3 SC and Article 7 of the ECRML shapes the legal status of Catalan in the eastern area of the Autonomous Community of Aragon. The status of the Catalan language – and Aragonese, as the two languages traditionally spoken in Aragon –, has evolved from the ambiguous Statute of Autonomy of Aragon of 1983 (SAA), through the 1996 amendment that foresees ‘their teaching and the speakers’ right to use them in the manner established by a law of the Cortes (Regional Assembly)’. A new 2007 SAA, even though it fails again to explicitly mention Catalan and Aragonese, as ‘the languages and linguistic forms native to Aragon’, or as ‘one of the most important expressions of the historical and cultural heritage of Aragon’ in Article 7, reinforces the previous mandate to legislators by stating:

‘A law of the Cortes of Aragon shall establish the regions of the predominant use of the languages and forms native to Aragon; it shall regulate the legal provisions, the speakers’ rights of use in those regions; it shall promote the protection, recovery, teaching, promotion and diffusion of the linguistic heritage of Aragon, and it shall favour, in the regions of predominant use, the use of local languages in the relations between Aragonese citizens and public administrations’ (Art. 7.2 SAA, cf. Vernet/Pons 2011, 65)

For many years, and in spite of constant social and political demands (starting with the *Declaration of Mequinensa*, signed in 1984 by mayors and councillors of seventeen municipalities in the Catalan-speaking region), the absence of a legal framework of statutory provisions deprived Catalan speakers of any legal guarantee of their proclaimed rights. Only the teaching of the language was granted – through collaboration with the education system of Catalonia – and voluntarily adopted by a broad majority of pupils in the Catalan-speaking municipalities. A fragmentary form of regulation emerged from the Aragonese Law 1/1999 of 24 February on inheritance, which recognizes the validity of wills drafted in Catalan, and Law 3/1999 of 10 March on the cultural heritage of Aragon, but this development remained far from the efforts made on behalf of Spanish under the ECRML, as denounced by the Council of Europe in the reports adopted in 2005 and 2008.

Law 10/2009 of 22 December on the use, protection and promotion of languages native to Aragon firstly recognizes a series of rights for speakers in primary and secondary education and in relations with public administration, in the ‘areas of predominant historical use’ or ‘mixed areas’ where Catalan is used, and also allows the use of Catalan place names. The law of 2009 faced some criticism for not consistently improving the institutional and public usage of Catalan. The legislative provision has not translated into public action favouring the use of Catalan, which is currently only present in place names, signposting and certain public and private documents (Vernet/Pons 2011).

However, following a political change in the autonomous institutions (the substitution of the Socialist Party by the Popular Party in the regional government) a new

Law 3/2013 from May on the use, protection and promotion of languages and linguistic varieties of Aragon (LLA) repealed the 2009 law. Despite exhibiting a similar structure, the new LLA, which includes mandates for safeguarding and promoting the languages and linguistic varieties of Aragon (Arts. 9 to 11) and the possibility of their use in the oral debates of local councils (Art. 20), in personal names (Art. 23) and in the media (Art. 24), also substitutes recognition of Catalan – and Aragonese – as protected *languages* with the acronyms of the area of their use (*LAPAO* for Catalan) and revokes previous guarantees of linguistic rights. The Constitutional Court Judgment 56/2017 of 17 March refused the appeal of various Spanish deputies against the step back derived from the LLA, declining to consider its compatibility with the ECRML. As a result of this inconsistent legislation, the language policies related to Catalan and any financial support it receives remain highly dependent on the political will of the current Autonomous Government. Some significant advances have been observed recently by the establishment, in 2015, of a General Directorate for Language Policy, including the adoption of linguistic curricula (Art. 13 LLA), the creation of the advisory committee on toponymy (Art. 22 LLA) and the adoption of the statutes of the Aragonese Academy of the Language (Art. 7 LLA). The declaration of the areas of use of the languages and linguistic varieties of Aragon (Art. 6) is still pending.

8 The Community of Valencia

The Statute of Autonomy of Valencia (hereafter, SACV) and the Valencian Law 4/1983 of 18 April for the use and teaching of Valencian (LUTV) established and developed the official status of Catalan, referred to as ‘Valencian’ in this Autonomous Community. Article 7 of the SACV of 1982 already recognizes Valencian and Castilian as the two official languages and their associated linguistic rights, in addition to other provisions inspired by the aim of *normalization*. After the reform of the SACV in 2006, the new Article 6 contains a definition of Valencian as ‘the own language of Valencia’ (“la llengua pròpia de València”), even if similar consequences to other linguistic legislations do not arise from this concept (Ochoa 2006). The legal denomination of the language (‘Valencian’) – with a sense of identity –, is not contrary to recognizing the unity of the Catalan language (CC Judgment 75/1997 of 21 April and other ordinary judgments that recognize the validity of Catalan proficiency certificates to assess linguistic knowledge for Valencian teachers, before the recent Valencian Order 7/2017 of 2 March, which approves Catalan certificates from other territories).

Article 6.2 SACV regulates the official status of Valencian and Castilian and recognizes that ‘everyone has the right to know and use them, and to receive education in and of the Valencian language’, even though such rights are partially subject to the establishment, by law, of the application criteria for the local language in public administration and in education (Article 6.6 of the SACV). Some forms of territorial modulation of this official status is foreseen in traditionally Castilian-speaking terri-

tories (Art. 6.7 SACV), as a special feature of this regional linguistic model, which includes the distinction between municipalities of Valencian or Castilian predominance (Arts. 35 and 36 LUTV). According to Article 37 LUTV, this distinction does not prevent public action to support the linguistic right to know and to use Valencian (Art. 1 LUTV).

The LUTV is generally not considered to be compliant with normalization laws, for it doesn't explicitly aim at normalizing the local languages use in different sectors (Vernet/Pons 2011). The content of the law is mainly addressed at developing the double official status of both languages, and many of its precepts establish a parallelism between Valencian and Castilian: for example, in administrative proceedings (Art. 11); courts of justice (Art. 12); public documents (Art. 13) or education (Art. 19). Nevertheless, the preamble of the LUTV refers to the commitment of the Valencian authorities to defending the cultural heritage of the Autonomous Community and 'to restore our language to the category and the domains it deserves, bringing an end to the situation of neglect and deterioration in which it finds itself'. The development of this mandate implies declaring that Valencian, as the language of the Community, 'is also the language' of local and autonomous public administrations (Art. 7), that is, declaring its special position/status as a language to be promoted in cultural sectors and social media (Art. 25 and 27 LUTV). By contrast, other rulings have an effect in favour of Castilian, even if they aim at correcting such an inequality (Art. 18 and 24, concerning education) or mitigating it (Art. 29 LUTV, concerning the promotion of the knowledge of Valencian among civil servants), occasionally with insufficient measures, if they are not generalized (for example, Art. 30.3 LUTV concerning the linguistic requirements for certain posts in the local or regional administrations).

In practice, compliance with the legal objective of promoting Valencian established in 1982 by the LUTV was only partially fulfilled. In fact, this law was initially approved by the Socialist Party, but during the long-term governments of the Popular Party a certain level of neglect and a conflicting vision of the language question (the Valencian Academy of the Language was created in 2000 in order to reach an agreement on standardization) did little to help promote Valencian. The decision to close the public media in Valencian, adopted by the Autonomous Government in 2013, represents a critical moment since, added to the political opposition to maintaining reciprocity regarding the access to Catalan-medium broadcasting from other Autonomous Communities, there was no television or radio content available to the public in Valencian anymore (a situation in breach of Art. 11 ECRML). However, the leftist majority resulting from the elections in 2015 has begun to pass new laws in a number of areas (Esteve 2018). These areas include education, with measures to reinforce the models that offer Valencian as a language of instruction (Law 4/2018 of 21 February on plurilingualism in the Valencian education system); public administration, with the establishment of Valencian as standard language (Valencian Decree 61/2017 of 12 May); and reinforced legal guarantees for language rights (Valencian Decree 187/2017 of 24 November, creating the Office for Linguistic Rights). A signifi-

cant number of place names were bilingualized or were adopted only in their Valencian form (e.g. *Castelló de la Plana*) during the latter period (Decree 69/2017 of 2 June). Finally, a major step was taken in 2018 with the restoration of Valencian public radio and television.

9 The Balearic Islands

The Statute of Autonomy of the Balearic Islands (hereafter, SABI) and the Balearic Law 3/1986 of 29 April for the linguistic normalization of the Balearic Islands (LLNBI) are the basic norms that regulate the Catalan language in this region. The SABI of 1983, similar to most *first generation* Statutes, includes a synthetic regulation of the official status of Catalan, largely maintained and expanded by its reform in 2007 (Colom 2009). The current Article 4 SABI defines Catalan as the ‘native language’ of the Community, in the context of a double official status of Catalan and Castilian, which implies that everybody has the right to know and use Catalan. It also includes a *normalization* mandate, according to which ‘the institutions of the Balearic Islands shall guarantee the normalized and official use of both languages, take the necessary measures to ensure proficiency in these languages and create the conditions to allow the full equality of both languages with regard to the rights of the citizens of the Balearic Islands’ (Art. 4.3 SABI). As for sectoral provisions, Article 35 SABI refers to the teaching of Catalan, including an explicit mention of the ‘insular forms of Catalan, of Majorca, Minorca, Eivissa and Formentera, without prejudice to the unity of the language’, while Articles 97 and 99 SABI see knowledge of Catalan as a preferred skill to access positions such as judge, notary and registrar.

In the Balearic Islands, the legislation on Catalan has been inspired by the aim of normalization. The LLNBI, adopted by all parliamentary parties present in 1986, proposes a sector-based objective ‘to implement the progressive and “normal” use of Catalan’ in the official and administrative domains, to guarantee the knowledge and progressive use of Catalan as a teaching language, to promote the use of Catalan in all domains of social communication, and to increase social awareness of the importance that all citizens know and use Catalan (Art. 1.2 LNBI). This legal text sets out the language rights and their protection by the courts, as well as the regulation of the Catalan language in public administration, education and the media, among others (Vernet/Pons 2011).

The basic provisions of the LNBI have been developed by sectoral legislation, which has been redrafted – as was the LNBI, partially abrogated by Law 9/2012 of 19 July, passed by the Popular Party majority, and restored by Law 1/2016 of 3 February by the new leftist majority – due to political changes in the Autonomous Government (Ballester 2016). In education, earlier advances followed from the transfer of competences in education by the state in 1994, establishing a conjunction model aimed at ensuring Catalan as language of instruction in 50 % of the subjects (in this

sector, the *trilingual model* designed in 2013 by the Popular Party failed following a major social contestation). In public administration, the standard use of Catalan in regional and local institutions, with corresponding linguistic requirements for civil servants, has also been restored by Law 4/2016 of 6 April (CC Judgment 165/2013 endorsed the step back introduced in this area by Law 9/2012). The status of the Catalan language as ‘the own language’ of public media permitted broadcasting in Catalan, and this offer is reinforced through the reception of public channels of other Catalan territories, such as Catalonia.

As far as the institutional promotion of Catalan is concerned, the situation in the socio-economic sector is highly influenced by immigration of Castilian speakers from the peninsula and the presence of German and English residents and tourists. Public efforts have been made to guarantee the linguistic rights of consumers as well as health care users, as stated by ECRML (Article 11), in the face of some resistance from the professional and business sectors. In practice, such linguistic influences make the recovery of the local language very difficult, particularly in the capital (Palma). The Institute of Balearic Studies carries out campaigns for the promotion of the language and culture of the Balearic Islands. Through its participation as a member of the Ramon Llull Institute the Autonomous Community of the Balearic Islands also contributes to the diffusion of the Catalan language and culture abroad.

10 Catalonia

The Statute of Autonomy of 2006 (SAC) and the Law 1/1998 of 1 January on language policy (LLP) synthesize the evolution of the legislation on language in Catalonia. Initially, the SAC of 1979 declared Catalan ‘the own language’ of Catalonia’ (“llengua pròpia de Catalunya”), establishing the joint official status of Catalan and Castilian, as well as initiating the process of linguistic normalization (according to Article 3.3 SAC, ‘the Autonomous Government shall guarantee the normal and official use of both languages, take the necessary measures to ensure their knowledge, and create the conditions that permit their full equality with regard to the rights and obligations of the citizens of Catalonia’). The first normalization act, the Law of 1983 for linguistic normalization in Catalonia, states that ‘the reestablishment of Catalan to its rightful place as a local language of Catalonia is an irrefutable right and obligation of the Catalan people’, regulates its use in public administration, education and public media and promotes its use in other areas. The LLP goes a step further, stating in its preamble that it seeks ‘to consolidate the process promoted by the Law of linguistic normalization [...] with the aim at advancing the generalization of the full knowledge and normal use of Catalan’. This law reinforced the regional control of the established linguistic system and advanced into sectors previously not regulated and where Catalan had not progressed as much, such as in legal proceedings (Art. 13 LLP), civil and commercial documents (Art. 15), public registers (Art. 17), the cultural industries

(Art. 28) and public relations in companies and establishments involved in the sale of products and services (Art. 32). This law is also important because it describes the significance and legal effects of the notions of ‘own language’ and ‘official language’ (Arts. 2 and 3 LLP) and lists the language rights of citizens (Art. 4) (Pons 2006; Vernet/Pons 2011).

The reform of the statute in 2006, which had as its main goal the improvement of Catalonia’s self-government, including the linguistic regulation, must be read primarily as a consolidation of the previous legal guarantees. The most significant provisions include the proclamation of the right and duty of citizens to know Catalan (Art. 6.2 SAC), by establishing a parallel with the official status of Castilian ex Article 3.1 SC; the definition of Catalan, as the ‘own language’ of Catalonia, as ‘the language of common and preferential use in the bodies of public administration and in the public media of Catalonia, and (is) also the language of normal use for teaching and learning in the education system’; some sector-based linguistic rights related to public administration, education and the socio-economic sector (Arts. 32–36 SAC), as well as the principle that members of the judiciary and other administrative staff employed by the state authorities in Catalonia, in addition to notaries and registrars, must have adequate and sufficient knowledge of Catalan (Art. 33.3, 33.4, 147 and 102 SAC). At the same time, the statute of 2006 includes several requirements with respect to the protection, promotion and projection of the Catalan language in the central institutions of the autonomous government, the state, the European Union and UNESCO (Arts. 6.3, 33.5 and 50.7 SAC).

Four years after the SAC was ratified in a referendum by the population of Catalonia, CC Judgment 31/2010 of 28 June, following the appeal from legislators of the Popular Party, largely endorsed the constitutionality of this legal-linguistic regime (with the exception of the ‘preferential’ use of Catalan foreseen in Art. 6.1). However, some interpretations of the CC have intensified hermeneutical debates and inspired subsequent rulings of both this court and ordinary jurisdiction over the contents of sectoral laws relating to the Catalan language. Consequently, all language laws adopted by the Catalan Parliament that developed the SAC, in areas such as the public media, immigration, consumer rights and cinema, have been challenged by a constitutional appeal (see, respectively, CC Judgments 86, 87, 88 and 89/2017 of 4 July). In the field of education, the consequences of the interpretative declaration of the CC of Castilian as a medium of instruction – as Catalan is according to Articles 6 and 35 SAC –, are observable in judicial decisions that require Castilian be used 25 per cent of the time if requested by at least one pupil in a class group (CC Judgment 14/2018 of 20 February denies the state the power to oblige the Generalitat to subsidise private education for students who ask for education in Castilian).

In CC Judgement 31/2010 it is evident that the devaluation suffered by the Statute of Autonomy is intended to restrict the region’s linguistic autonomy (Art. 143 SAC has established an exclusive competence of the Generalitat concerning Catalan). Simultaneously, the CC emphasizes the supremacy of the official status of Castilian – as a

reflection of the unity of the state – and weakens the official status of the Catalan language through a number of different arguments: the effects of the ‘general duty’ for all citizens to know (only) Castilian, the advances in the normalization of Catalan as a counterargument to positive discrimination measures and the very strict limits derived from other state norms for differentiated regulations for Catalan. All this demonstrates the significance of Judgment 31/2010 of the CC in declaring the notion of ‘preferential’ treatment for the ‘own language’ unconstitutional and imposing a ‘strict equality’ of both official languages, which is not contradictory – in the view of the CC – with the prevailing position of Castilian in the Catalan-speaking territories.

11 Final considerations

The analysis of the evolution of the regulation of the Catalan language in various regions shows how politics has transcended mere linguistic facts. Granting official status to a given language is a crucial political decision, with significant implications for the linguistic rights of speakers and language policies aimed at protecting and promoting this language. However, as noted above, while this official status obviously matters, it is not sufficient for the effective recognition and development of language usage (Pons 2017). Applying these general considerations to the Catalan language, it would seem adequate to encourage an improvement of its status in those territories where language rights are not granted and there are no clear criteria for its treatment when it comes to the distribution of public resources. Conversely, with respect to other territories where this official status is formally recognized, the effective normalization of the Catalan language involves various elements of the legal framework that delimit the scope of its official status and official use in some areas. In that sense, we can argue that language normalization can only be partially achieved through the action of the regional or autonomous authorities and cannot succeed if the government of the state does not support this official status or tries to minimize legal guarantees around it, subjecting any positive measure to conditional permission. For example, this was the case with the aforementioned CCJ 31/2010, which established that Catalan legislators may adopt necessary and proportionate language policy measures that favour Catalan ‘in the case’ of a historical misbalance of one of the two existing official languages.

Particularly in Catalonia, the advances in the process of linguistic normalization and the attempts to guarantee language rights and prerogatives related to Catalan, as the own language of this territory with higher-rank norms, has led to a negative reaction from some political parties and state authorities. Legislative evolution in Catalonia cannot be isolated from a general constitutional framework which privileges the language rights and duties of Castilian not only in its traditional (legally monolingual) areas, but also in (legally bilingual) Catalan areas, and from the language policy of the Spanish state – sometimes implicit in other policies – oriented towards reinforcing the status of the Castilian language with its pre-existing advantage in

demolinguistic terms. In that sense, it is not only the central institutions of the state that remain monolingual (the Constitutional Courts, the state government, Parliament –with the small exception of some activities of the Senate –, Ombudsman and Higher Courts), but the state also uses its authority to impose the required or exclusive use of Castilian in other areas such as labelling, administrative procedures and platforms, foreign affairs, etc. One example is the judiciary, subject to state authority, in the Autonomous Communities where this double official status has been sanctioned since the 1980s and where Castilian continues to be defined as ‘the language of Courts’ (Organic Law 6/1985 of the Judiciary Power, Art. 231) and there are no legal guarantees regarding the knowledge of Catalan by judges and magistrates, a situation that is contrary to the obligations assumed by Spain under Article 9 of the ECRML.

The political conflict triggered by the CC Judgment 31/2010 – with long-term historical roots – can be interpreted as a reaction to the perceived threat to the relevant social consensus built on the basis of the SAC of 1979, which includes language law and language policies related to Catalan which oppose assimilation policies promoted by the state (for example, the Organic Law 8/2013 on the improvement of the quality of education, which undermines the basis of the Catalan model of conjunction aimed at promoting bilingualism among the population). The secessionist process that had its pivotal moment with the referendum on self-determination of 1 October 2017, and the subsequent repression from state authorities, reflect the need for recognition of the political personality of Catalonia and show the limits of the current State model to face this demand. While the linguistic element has not occupied a central position in these demands (which are mainly based on democratic and economic grounds), widespread suspicions regarding the Catalan-language policies expressed from unionist positions points to its importance in the historical continuity of a community that seeks to preserve its identity, whilst at the same time remaining an open society.

12 Bibliography

- Arzoz, Xabier (2009), *Lenguas y modalidades lingüísticas en la Constitución Española: ¿dos regímenes jurídicos diferenciados?*, in: Antoni Milian (ed.), *El plurilingüisme a la Constitució Espanyola*, Barcelona, Institut d'Estudis Autònoms, 61–122.
- Ballester, Maria (2016), *Una nova empenta per a l'ús social de la llengua catalana a les Illes Balears. L'avanç inestable, i no sempre lineal, de la normalització lingüística*, *Revista de Llengua i Dret/ Journal of Language and Law* 66, 85–104.
- Bastardas, Albert/Boix, Emili/Torrens, Rosa M. (coord.) (2018), *El català, llengua mitjana d'Europa. Multilingüisme, globalització i sostenibilitat lingüística*, Barcelona, Octaedro.
- Cabellos, Miguel Ángel (2008), *La competència en matèria de llengua pròpia en el nou Estatut*, *Revista de Llengua i Dret/ Journal of Language and Law* 49, 69–96.
- Caria, Marco (2014), *Alghero-l'Alguer o il catalano d'Italia*, *Bollettino dell'Atlante Linguistico Italiano* 38, 75–90.
- Caria, Rafael (2006), *El català a l'Alguer: Apunts per a un llibre blanc*, *Revista de Llengua i Dret/ Journal of Language and Law* 46, 29–102.

- Colom, Bartomeu (2009), *La reforma de l'Estatut d'autonomia de les Illes Balears de 2007 i la llengua catalana*, Revista de Llengua i Dret/Journal of Language and Law 52, 275–326.
- Council of Europe (2005), *Report of the Committee of Experts on the Application of the Charter in Spain*, 1st monitoring cycle, <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806dba66> (last accessed: 10.07.2018).
- Council of Europe (2008), *Report of the Committee of Experts on the Application of the Charter in Spain*, 2nd monitoring cycle, <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806dba68> (last accessed: 10.07.2018).
- Council of Europe (2012), *Report of the Committee of Experts on the Application of the Charter in Spain*, 3th monitoring cycle, <https://rm.coe.int/CoERMPublicCommonSearchServices/DisplayDCTMContent?documentId=09000016806dba6b> (last accessed: 10.07.2018).
- Council of Europe (2016), *Report of the Committee of Experts on the Application of the Charter in Spain*, 4th monitoring cycle, <https://rm.coe.int/16806f0658> (last accessed: 10.07.2018).
- Esteve, Alfons (2018), *Drets i deures lingüístics del Govern del Botànic*, Blog Revista de Llengua i Dret, 10 mai, <http://eapc-rld.blog.gencat.cat/2018/05/10/drets-i-deures-linguistics-del-govern-del-botanic-alfons-esteve/> (last accessed: 22.07.2018).
- Jou, Lluís (2011), *La STC 31/2010 reinterpretada. Legislació lingüística, realitat social i política*, Revista d'Estudis Autònoms i Federals 12, 153–191.
- Kraus, Peter (2018), *Between Minority Protection and Linguistic Sovereignty*, Revista de Llengua i Dret/Journal of Language and Law 69, 6–17.
- Lasagabaster, Iñaki (2017), *Derecho público de Heuskal Herria*, Bilbao, Lete Argitaletxea.
- Marí, Isidor (2011), *Policies Governing the Use of Languages in Relations Between the Authorities and the Public*, in: Miquel Strubell/Emili Boix (edd.), *Democratic Policies for Language Revitalisation: the Case of Catalan*, Basingstoke, Palgrave Macmillan, 84–118.
- Milian, Antoni (2009), *El marc constitucional espanyol relatiu al plurilingüisme: model reeixit o regulació obsoleta?*, in: Antoni Milian (ed.), *El plurilingüisme a la Constitució Espanyola*, Barcelona, Institut d'Estudis Autònoms, 209–258.
- Milian, Antoni (ed.) (2012), *Language Law and Legal Challenges in Medium-Sized Language Communities. A Comparative Perspective*, Barcelona, Institut d'Estudis Autònoms.
- Milian, Antoni (2016), *Más sobre derechos lingüísticos. Reflexiones sobre los límites constitucionales y su interpretación por el Tribunal Constitucional*, València, Institut d'Estudis Autònoms/Tirant lo Blanch.
- Navarro Sánchez, Ángel Custodio (2018a), *Sobre la protecció i reconeixement jurídics de l'expressió lingüística del català més desconeixuda i abandonada de totes: la del Carxe, a la Regió de Múrcia (I)*, Revista de Llengua i Dret/Journal of Language and Law 69, 176–209.
- Navarro Sánchez, Ángel Custodio (2018b), *Sobre la protecció i reconeixement jurídics de l'expressió lingüística del català més desconeixuda i abandonada de totes: la del Carxe, a la Regió de Múrcia (II)*, Revista de Llengua i Dret/Journal of Language and Law 70, 86–102.
- Nogueira, Alba/Ruiz, Eduardo J./Urrutia, Iñigo (edd.) (2012), *Shaping Language Rights. Commentary on the European Charter for Regional or Minority Languages in Light of the Committee of Experts' Evaluation*, Strasbourg, Council of Europe.
- Observatori de la Llengua Catalana (2015), *VII Informe sobre la situació de la llengua catalana, Barcelona*, <http://blogs.iec.cat/cruscat/wp-content/uploads/sites/15/2011/11/Informe-2014.pdf> (last accessed: 15.07.2018).
- Ochoa, Josep (2006) *Estatuto jurídico del valenciano*, in: José Manuel Pérez (ed.), *Estudios sobre el estatuto jurídico de las lenguas en España*, Barcelona, Atelier, 349–385.
- Patten, Alan/Kymlicka, Will (2003), *Introduction. Language Rights and Political Theory: Context, Issues, and Approaches*, in: Will Kymlicka/Alan Patten (edd.), *Language Rights and Political Theory*, Oxford, Oxford University Press, 1–51.

- Paz, Moria (2013), *The Failed Promise of Language Rights: A Critique of the International Language Rights Regime*, Harvard International Law Review 54, 157–218.
- Pons, Eva (2006), *Los derechos lingüísticos en el marco internacional y europeo*, in: José Manuel Pérez Fernández (ed.), *Estudios sobre el estatuto jurídico de las lenguas en España*, Barcelona, Atelier, 65–104.
- Pons, Eva (2011), *El dret lingüístic en l'àmbit català*, Treballs de Sociolingüística Catalana 21, 115–124.
- Pons, Eva (2013), *The Effects of Constitutional Court Ruling 31/2010 Dated 28 June 2010 on the Linguistic Regime of the Statute of Catalonia*, Catalan Social Sciences Review 3, 67–92.
- Pons, Eva (2015), *L'oficialitat lingüística. Declaracions constitucionals i implicacions jurídiques i pràctiques*, Barcelona, Generalitat de Catalunya.
- Pons, Eva (2017), *Presentation. Language and Status: Does Official Status Matter?*, Revista de Llengua i Dret/Journal of Language and Law 67, 1–6.
- Segura, Lluís (2011), *Les llengües oficials en la doctrina recent del Tribunal Constitucional*, Revista de Llengua i Dret/Journal of Language and Law 56, 83–113.
- Toso, Fiorenzo (2008), *Le minoranze linguistiche in Italia*, Bologna, il Mulino.
- Vernet, Jaume (2007), *El pluralismo lingüístico*, in: Ramon Punset/Jaume Vernet, *Lenguas y Constitución*, Madrid, Iustel, 17–60.
- Vernet, Jaume/Pons, Eva (2011), *The Legal Systems of the Catalan Language*, in: Emili Boix/Miquel Strubell (edd.), *Democratic Policies for Language Revitalisation: the Case of Catalan*, Basingstoke, Palgrave Macmillan, 57–83.
- Vernet, Jaume, et al. (2003), *Dret lingüístic*, Valls, Cossetània.
- Zabaleta, Eneritz (2015), *Principios constitucionales sobre las lenguas en Francia*, Revista de Llengua i Dret/Journal of Language and Law 63, 93–112.

F. Xavier Vila

20 Teaching and Learning of Catalan

Abstract: Educational policies have been a battlefield for language policies in the Catalan-speaking territories since the 18th century, when these territories lost their right to self-govern. Unsurprisingly, language-in-education policies have also been the spearhead of Catalan language recovery, especially since the late 1970s when the opportunity to develop autonomous school language models arose. This chapter reviews the different language-in-education models developed in each territory for primary and secondary education, higher education and other forms of adult language education.

Keywords: language-in-education policy, Catalan conjunction model, language immersion, language proficiency, bilingual education

1 Primary and secondary education

1.1 Language-in-education policies until the late 1970s

‘We don’t know, and have no way of knowing when the teaching of Castilian in Catalanian schools started [...]. In primary schools, it is reasonable to assume that [...] it did not start until well into the 18th century. But it is also reasonable to think that some teachers, especially in Barcelona, and also in some other cities and towns, began to teach it before: those who were able to write it must have learnt it somewhere’ (Marfany 2001, 405).¹

There is little doubt among historians that until well into the 18th century, with the exception of some aristocratic institutions, Catalan was the medium of instruction in schools in the Catalan-speaking territory for those who had the means to have some sort of formal education (Torres Sans 2011, 83). This was not a majority of the population, but nevertheless included a sizeable part of the (mostly male) urban middle classes, artisans and even landed peasants (Marfany 2001; Montoya Abad 2012; Torres Sans 2011). Latin, of course, was learnt in grammar schools by those who were expected to go to university, or in seminars for those preparing for an ecclesiastic

¹ “No sabem, ni tenim manera de saber, quan es va començar a ensenyar el castellà a les escoles de Catalunya [...]. A les escoles de primeres lletres, sembla raonable de suposar [...] que no va ser fins molt entrat el XVIII. Però també ho és de pensar que alguns mestres, a Barcelona sobretot, però també en altres ciutats i viles, degueren començar a fer-ho abans: en alguns llocs havien d’haver-lo après els qui sabien escriure’l” (Marfany 2001, 405).

career. Castilian, a language used for high political, literary and religious purposes, seems to have occupied a discreet position at schools until the 18th century.

The first Catalan-speaking territory where Catalan was replaced by the language of the monarchy was Northern Catalonia. The first schools designed to teach (in) French in this territory were created in 1672. South of the Pyrenees, the Nueva Planta decrees abolished Catalan institutions and imposed Castilian as the language of administration. However, in spite of demands to this effect, the language of instruction at schools was not immediately targeted (Ferrer i Gironès 1985; Prats 1995, 325). It was not until 1768 that a Real Cédula, i.e., a royal decree, imposed the provision of primary schooling in Castilian. Historians disagree about the actual pace at which this decree was implemented, but available data suggest that it was only after the imposition of compulsory instruction by the Education Law of Moyano (1851) that Catalan was actually banned as a means of instruction and even as a subject of study (Marfany 2001, 413; Pueyo 1996).

The reintroduction of Catalan in education became one of the crucial demands of the growing Catalanist movement (Anguera 1997, 209–212), which in 1898 created the Association to Protect Catalan Education. Politicians as well as renowned pedagogues pledged for the normalization of Catalan at school, and this was a goal that both the Mancomunitat and, especially the Generalitat, started to implement (Monés 1984). However, the defeat in the 1936–1939 Civil War meant that Catalan was again banished from schools. The end of Franco's dictatorship and the establishment of the 'State of autonomies' system in Spain after 1978 (19 Language Law and Language Policies), opened a window of opportunity for a transformation of the language policies that varied greatly among the Catalan-speaking territories (Vila 2008; 2011).

1.2 Catalonia

In Catalonia, popular demands to reintroduce Catalan in schools were taken up by the local authorities. The process was relatively rapid: at the end of the 1970s only a handful of schools taught in Catalan and the teaching of the language was not yet common in compulsory education; two decades later most schools were using Catalan as their main language of instruction. A number of elements were crucial to the success of this process (Canal Santos/Marco Escabosa 2016). To start with, even if a large percentage of the population had settled in Catalonia in the previous decades (18 Language Demography), mastery of the language was viewed by natives and migrants alike as an instrument of social integration and progress. Secondly, the school system was designed as an additive bilingual model which expected students to reach the same levels of proficiency in Castilian as schools with Castilian as the main language of instruction. Thirdly, the Catalan school model was based on an integrationist ideology which rejected the separation of students (and teachers) according to first language – for this reason it is technically referred to as the Catalan

conjunction model (CCM). A Catalan language immersion program was implemented in classes where more than 70 % of the pupils were non-Catalan L1 speakers. The CCM helped to forge a strategic alliance between all political parties with a professed Catalanist sensibility, i.e., at least 80 % of the members of the Catalan parliament.

By the end of the 1990s, the CCM had a mixed balance sheet. On the one hand, schools in Catalonia were effectively providing high levels of oral and written proficiency in Catalan and Castilian to a majority of students, whilst simultaneously avoiding the risk of social division according to first language (Woolard 1997). On the other hand, the model was not increasing the use of Catalan among students (Vila 2014; Vila/Vial i Rius 2005). Besides that, the CCM had become a frequent target for the Spanish conservative press and politicians who argued that Castilian should be given a more prominent role.

The ‘new immigrations’ (“noves immigracions”) in the 2000s represented a turning point (718 Language Demography). The arrival of huge numbers of children with different linguistic repertoires challenged the schools’ routines. A new ‘immersion model’ (“model d’immersió lingüística”) was developed, under which welcoming classes and ‘environmental plans’ (“plans educatius d’entorn”) were set up (Arnau/Vila 2013, 12–18; Vila i Mendiburu 2012).

The 2010s saw a significant transformation of language-in-education policies in Catalonia. The end of massive immigration lessened the demand to manage linguistic diversity in the classroom. Simultaneously, the court ruling on the 2006 Statute of Autonomy of Catalonia (719 Language Law and Language Policies) prompted some judicial conflicts regarding the use of Castilian as a means of instruction (Mayans/Tort/Areny 2016). Besides that, the real change on the ground was the growing demand for improved English proficiency, which has encouraged the spread of CLIL courses, especially in the private sector. In 2018, in the midst of considerable social debate surrounding the need to reform education, the Catalan Ministry of Education published a document outlining the new *Catalan plurilingual and intercultural model* (Subdirecció General de Llengua i Plurilingüisme 2018). The model is based on the concepts of plurilingual and pluricultural competence developed in the European Framework, and aims at fostering high levels of proficiency in Catalan, Castilian and one or two foreign languages – plus Occitan, Catalan Sign Language and heritage languages wherever relevant –, and at promoting the use of Catalan.

1.3 The Valencian Country

The Law of Use and Teaching of Valencian passed by the Valencian Parliament in 1983 was less ambitious than its Catalanian equivalent (719 Language Law and Language Policies). In educational terms, it led to the creation of different language programmes, which were, in primary education: (a) the *Progressive Incorporation Programme* (PIP), in which Castilian was the main language of instruction and

Catalan a compulsory subject and the means of instruction for some subjects; (b) the *Teaching in Valencian Programme* (PEV), with Catalan as the means of instruction with Castilian as a compulsory subject and used as a means of instruction for some subjects; and (c) the *Language Immersion Programme* (PIL), aimed at students whose first language was not Catalan and in which teaching was carried out mainly in Valencian. In secondary education, PIL and PEV were merged. In Castilian-speaking areas, only a Castilian-language *Basic Programme* with optional teaching of Catalan was implemented. Although they made some progress over time, the number of Catalan-language programmes remained systematically below Castilian-language ones, to a large extent due to the lack of support from the conservative administrations in power between 1995 and 2015. In 2011, for instance, only 29 % of all pupils were enrolled in Catalan-language schools, whereas in 2014 the figure had only risen to 30 % (Allioli 2016, 4). The evolution of language proficiency (↗18 Language Demography) made it clear that only PEV and PIL resulted in bilingualism and biliteracy.

The late 1990s saw the appearance of Enriched Bilingual Education Programmes (PEBE), which introduced English as a means of instruction in combination with one of the existing programmes (PIL, PEV or PIP). In 2012, programmes were reconfigured into two plurilingual programmes in which the three languages were supposed to be used: PPEV, with Valencian as the main language of instruction, and PPEC, with Castilian as the main vehicular language. This change meant the abolition of immersion programmes and an increase in the use of English, but did not improve the position of Catalan (Allioli 2016).

The political transition from a conservative government to the progressive authorities in 2015 opened the door to the overall transformation of Valencian language policy (↗19 Language Law and Language Policies). The 9/2017 Decree of Plurilingualism replaced the existent programmes with a new Dynamic Plurilingual Education Programme designed to contribute to the development of ‘a multilingual and multicultural Valencian society, with the special protection of Valencian as a language of integration and cohesion’.² According to this decree, all schools were expected to use Catalan, Castilian and English as a means of instruction, but each school was allowed to design its own specific Language Plan and place itself on a continuum of 6 categories ranging from Basic I (Castilian as the predominant means of instruction) to Advanced II (Catalan as the predominant means of instruction with a significant presence of English). In this scheme, students would be granted EFLT official recognition of their proficiency level in Catalan and English in accordance with the level of plurilingualism adopted by their school. By March 2017, 54 % of infant or primary education centres had opted for an advanced level, 30 % for an intermediate level,

² Art. 8.1.j: “Col·laborar en la consecució d’una societat valenciana multilingüe i multicultural, amb especial protecció del valencià com a llengua d’integració i cohesió.”

and only 16 % for a basic level. The evolution from 2016, when only 34 % had opted for predominantly Catalan-language schools, was substantial (Gabinet de Comunicació 2017). In spite of this support, opposition to the decree from the conservative parties blocked its implementation and forced the Valencian Government to present a new Law on Plurilingualism in the Educational System. The 4/2018 Law retained the basic philosophy of the decree except for official accreditation of the proficiency level in foreign languages and was passed in February 2018 by the Valencian Parliament.

1.4 The Balearic Islands

Changes in language-in-education policies in the Balearic Islands happened slowly. In 1986, the Language Normalization Act required that schoolchildren be able to use Catalan and Castilian correctly by the end of compulsory education, but the education system was in the hands of the central government until 1998 and few measures were adopted to encourage bilingualism. In fact, it was only in 1997 and after large-scale popular mobilization that the autonomous parliament, in the hands of the conservative PP, passed the so-called “Decree of Minimums”, requiring that at least half of all school subjects be taught in Catalan. The subsequent centre-leftist government (1999–2003) started to implement this decree, including a provision requiring all teachers to demonstrate knowledge of Catalan.

On the ground, the combination of the decree and the transfer of educational responsibilities to the island authorities has benefited the position of Catalan. By the academic year 2011–2012, a majority of pupils attended Catalan-medium schools: 73 % in kindergarten, 70.4 % in primary education, 64.9 % in compulsory secondary education (ESO), 69.7 % in Baccalaureat, and 69.1 % in vocational training (Spanish Ministry of Education, Culture and Sport, 2014 quoted by Mayans/Tort/Areny (2014, 115). Simultaneously, though, language-in-education policy has become a hot political issue. In 2003, the conservative PP was voted back into the autonomus government and passed a decree designed to encourage the use of foreign languages in teaching in principle and to allow for a reduction in the use of Catalan to only one third of teaching hours. These and other initiatives encountered strong opposition from the educational community and very few centres implemented them. In 2007, the new centre-leftist government revoked these measures and implemented some measures to promote trilingual education. The comeback of a radicalized PP in 2011 entailed a dramatic reduction in the presence of Catalan in all public spheres of life including administration, mass media, toponymy, and, of course, education (19 Language Law and Language Policies). The Integral Treatment of Languages (TIL) Project adopted by the conservative government explicitly sought to minimize the use of Catalan to one third of school time, a measure which worked in favour of the use of Castilian and, wherever possible, English. Nevertheless, the project encountered fierce opposition from the educational community – 79 % of parents opted to

have their children educated mainly in Catalan –, including the longest strike and largest demonstration in the islands' history (González 2013). The project was ultimately not implemented and, on the contrary, contributed decisively to the notorious defeat of the PP in the 2015 election (Mayans/Tort/Areny 2016). The leftist parties that won abolished the TIL, returning to the Decree of Minimums and announcing a new policy that would protect Catalan and encourage foreign language learning.

1.5 La Franja

Catalan is not an official language in la Franja in Aragon and, as a result, its presence in schools is very weak. In 1985, within the framework of the collaboration agreement with Spain's Ministry of Education and Science, schools in la Franja began to teach Catalan on a voluntary basis. Catalan classes began in the 1984–1985 academic year with 700 students and by 2004–2005, they were being taken by around 3,500. Additionally, a number of primary and secondary schools have some “bilingual” programmes, where Catalan is used for one subject (Mayans/Tort 2011).

The legislative changes restricting Catalan between 2013 and 2016 (719 Language Law and Language Policies) did not have any practical impact on the teaching and learning of the language. The majority of the school population follows the Catalan courses (2 to 4 hours a week). Those pupils that have followed these courses at least during compulsory secondary education and four years of primary education receive a B2 level accreditation in Catalan (Mayans/Tort/Areny 2016, 115). In real terms, these courses do not guarantee that non-native speakers acquire productive competence.

1.6 Andorra

The educational landscape in Andorra is quite complex as, in practical terms, several education systems coexist within its territory: the Andorran system, the French system, and the Spanish system.³ Most schools in the Spanish system (in Catalan *congregacional*, i.e., Catholic schools) use Catalan as their means of instruction, which leaves just the Spanish school and the Spanish Institute teaching entirely in Castilian. The reason for this complexity is that in the 19th and 20th centuries, Andorrans were essentially served by foreign education systems – the French and the Spanish – which granted them free education and access to their universities.

Massive immigration in the 1960s and 1970s changed the status quo, and Andorran authorities started to implement their own education policies. On the one hand, the authorities required the teaching of Catalan and Andorran history and culture in

³ <https://www.educacio.ad/sistemes-educatius>(last accessed: 20.03.2017).

all systems. On the other, a new Andorran education system was created in 1983–1984. This system was from inception that of a bilingual school, with Catalan as the main language of instruction and an important role attributed to French. In 2015–2016, the Andorran system catered for 40 % of pupils, whereas the French was attended by 33 % and the Spanish one 28 % (Mayans/Tort/Areny 2016, 99).

1.7 Northern Catalonia

Catalan finds itself in a very weak position in Northern Catalonia both in demolinguistic and legal terms (↗18 Language Demography and ↗19 Language Law and Language Policies). Having been completely banished from schools for three centuries, its position in the education system has improved over recent decades since the 1951 Deixonne Law (Loi n^o 51–46).

The supply of Catalan language courses in Northern Catalonian schools takes many different forms. In the first place, Catalan lessons are offered in a number of mainstream schools in different formats including basic initiation. Catalan language learning is also provided by a variety of bilingual schools which range from a very limited use of Catalan in a predominantly French-language model, to the equal use of French and Catalan. Finally, a partly-subsidised cooperative network of Catalan-language schools (La Bressola, i.e. ‘the cradle’) provides for over 1,000 pupils from pre-school (2 years old) to secondary education in 7 schools and 1 collège (ACN 2016). These schools use Catalan as their main means of instruction and encourage its interpersonal use (La Bressola 2007).

Whereas up to 70 % of parents wanted their children taught in Catalan according to an official survey presented in 2017 (Renyé 2017), the diversity of options makes it difficult to evaluate the actual impact of Catalan language education in Northern Catalonia. At the beginning of the 2017–2018 course, 14 % of schools of the Department offered a partial or total immersion programme which were followed by 19 % of pupils, (DSDEN 66 2018), and the total number of students with some sort of Catalan language instruction was around 17 % (Mayans/Tort/Areny 2016, 113)

1.8 L’Alguer

The social situation of Catalan in l’Alguer is very fragile (↗18 Language Demography), and the presence of the language in schools has experienced ups and downs. The 2008 economic crisis has been particularly damaging (Mayans/Tort/Areny 2016). The Palomba Project, set up in 1998 to teach Alguerese Catalan to primary school children at almost all local schools was discontinued in 2012–2013 due to financial problems. The trilingual (Catalan/Italian/English) “La Costura” private infant school which started in 2004–2005 with the support of the local city council as well as Catalonia’s

Department of Education, was also terminated in June 2014. Alternatively, Catalan language courses were resumed in 8 educational centres thanks to the Regional Law of 9 March 2015.

The legal situation of Catalan in l'Alguer changed significantly on 28 June 2018 when the Sardinian Regional Council passed the *Disciplina della politica linguistica regionale* or 'Law of regional language policy'. The new legislation foresees the promotion of Sardinian languages, including Catalan in l'Alguer, in public administration, mass media and other domains, and opened the door for Catalan-medium education (ACN 2018).

2 Higher education

Universities have operated in the Catalan-speaking territories since the Middle Ages, starting with the university of Lleida, founded in 1300, and followed by institutions in Perpinyà (1350), Barcelona (1450), Girona (1446) and Valencia (1499) amongst others. For centuries they used Latin as their main *lingua academica*. The annexation of Northern Catalonia by France in 1659 led to the adoption of French at Perpinyà University. The annexation of the rest of the language area by Castile after the War of Succession (1700–1714) brought about the dissolution of all Catalan universities and the creation of the University of Cervera that was controlled by the absolutist central authorities and was castilianized, as was its Valencian sister. It was only in 1837–1842 that Cervera was replaced by the new University of Barcelona, though universities continued to function exclusively in Castilian.

Towards the end of the 19th century, the growing Catalanist movement demanded that Catalan be reintroduced as a language of higher education. In 1903 a group of intellectuals and professors launched the *Estudis Universitaris Catalans* (Catalan University Studies) to teach linguistics, civil rights and other subjects in Catalan. Four years later, in 1907, the Institut d'Estudis Catalans (Institute of Catalan Studies) was created as the highest scientific academy in Catalonia. The elevation of Catalan to higher education was completed in 1933 when it became an official language of the University of Barcelona. The progress of Catalan as a means of instruction was rapid: during the academic year 1932–1933, the language of instruction was Catalan in 124 subjects and Castilian in 130, and the local language was already predominant in the Faculties of Medicine and Philosophy and Letters (Nadal Badal 1993, 148). The expansion of Catalan was nevertheless cut short by political conflicts and, eventually, the Republic's defeat during the Spanish Civil War (1936–1939).

The abolition of Catalan autonomy as a result of the war led to the immediate banishment of Catalan from higher education. During the postwar period, cultural resistance only managed to recover lost ground very slowly, such as through the clandestine activities of the IEC in the 1940s and courses in language and literature offered by the EUC in the 1950s. It was not until the late 1960s that some university

professors dared to give the first lectures in Catalan. Only three years after Franco's death (1975), though, the number of courses taught in Catalan had risen to 50 % in the three universities in Catalonia (Oliva/Vinent 1980). Since then, universities have been especially open to Catalan, particularly in Catalonia and, to a lesser extent, the Balearic Islands, which obtained its first university in 1978.

Regarding language policy, it should be kept in mind that universities in the Catalan territories have been, until the present, more or less public. Language policy at the universities in Catalonia and, to a large extent, the Balearic Islands, is organised based on several key principles (Pons i Parera 2015 Vila 2018):

1. *The officiality principle*: Catalan and Castilian enjoy official status, i.e., they can be used by any member of the university community, and all university members are supposed to understand both of them.
2. *The 'own language' principle*: Catalan is the language of preferential use in universities as administrative bodies and for educational activities.
3. *The linguistic normalisation principle*: Universities must take appropriate measures to ensure the use of Catalan in all spheres.
4. *The language conjunction principle*: Students should not be separated into centres or class groups on the basis of their habitual language of use.
5. *The internationalisation principle*: Universities can establish specific criteria for language use in activities related to international commitments.

In the Valencian Country, language policy in higher education is based on the development of Catalan language streams which are basically added to the majority Castilian-language streams.

On a personal level, for most of the university community in Catalonia and the Balearic Islands, the combination of these principles can be synthesized as the “emissive freedom plus obligation to understand” policy, which implies that all members of the university community are allowed to express themselves in either official language but are supposed to understand both of them, i.e., reciprocation (answering in Catalan/Castilian to someone who speaks Catalan/Castilian) is a matter of personal choice. Professors are in principle allowed to decide in which language they wish to teach their courses, and students can use either official language to interact with them.

This organisation of languages has so far proved to be rather peaceful and has allowed for the significant use of Catalan, which accounts for 77 % of undergraduate and 55.3 % of postgraduate courses in Catalonia during the academic year 2016–2017 (Departament de Cultura 2016, 35, 36). The system is working because the vast majority of students are proficient in both official languages and because Catalan is a Romance language, which means that students originating from other Romance-speaking countries – not only Spain, but also Italy, Portugal or France – find it relatively easy to follow courses in either language. Nevertheless, increased internationalisation has resulted in a growing number of non-Romance L1 speaking students

attending courses at Catalan universities. This has prompted a number of measures, including an increase in the number of courses taught in English – accounting for 20.4 % of postgraduate courses in the academic year 2016–2017 in Catalonia (Departament de Cultura 2016, 33) –, often also attended by local students who want to improve their command of the language. It has also led to the introduction of a sixth principle, the *linguistic transparency/security principle*, according to which the language of tuition, which is officially announced before students enroll in a course, cannot be renegotiated during the course. This principle was introduced to prevent students who had enrolled in courses advertised as being taught in Catalan from being able to demand teaching in Castilian or even in English by arguing that they were not proficient in the language of instruction. In the Valencian Country, on the contrary, Catalan is less widespread as a means of instruction in higher education, and usually included in Valencian streams (Aparici et al. 2011).

3 Adult language education

It was under the Francoist dictatorship, in the context of massive immigration of non-Catalan speakers and a language shift in some regions and social strata (718 Language Demography), that a number of different organisations were created to promote Catalan language teaching and learning. These institutions, such as Òmnium Cultural (1961–) in Catalonia and Obra Cultural Balear (1962–) in the Islands, were financed by popular subscription and private sponsors and had to work independently from public institutions. Their example encouraged the appearance of Acció Cultural del País Valencià (1978–) in the Valencian Country, Obra Cultural (1985–) in l'Alguer, as well as different local institutions that joined forces in the Iniciativa Cultural de la Franja (2003–).

After the fall of the dictatorship, public institutions started to implement adult Catalan language education programmes. Catalonia led the way in this process, and its impulse has benefited the rest of the territories. The Direcció General de Política Lingüística (DGPL) (1980–) boosted the resources in terms of courses and materials, which included, among others, the journal *Com ensenyar català als adults* ('How to teach Catalan to adults') (1982–1997) (Gabinet de Didàctica 1997), which was eventually succeeded by *Llengua i ús: revista tècnica de política lingüística* ('Language and use: technical journal of language policy').⁴ The DGPL also produced the multimedia Catalan language course *Digui, Digui* (1984–2007), which was replaced by the internet *Parla.cat* course in 2007. All of these initiatives were also possible thanks to the vigorous renewal in theoretical and applied approaches at the university departments

⁴ http://llengua.gencat.cat/ca/serveis/informacio_i_difusio/publicacions_en_linia/llengua_i_us/ (last accessed: 15.03.2017).

of Catalan Philology and Language Didactics, and the Institutes of Educational Sciences. The celebration of the first Symposium on Teaching Catalan to Non-Catalan Speakers in Vic in 1981, for instance, created a forum for face-to-face discussion that remains to this day.⁵

Adult language teaching in Catalonia took a significant step forward in 1989 when the different public institutions offering Catalan language courses targeted at the adult population outside of the universities coordinated their efforts in the Consortium for Language Normalization (CPNL).⁶ By its 25th anniversary in 2014, the 22 CPNL centres, which employed approximately 700 (full-time and part-time) workers, had given 62,325 language courses of various sorts, had had 1,535,965 registrations in these courses, and had organised 81,689 language partnerships between speakers and learners.⁷ Besides the CPNL, private language academies and NGOs also regularly offer tuition in the Catalan language.

A significant challenge for adult Catalan language teaching was that of adapting to the 'new immigrations' (↗18 Language Demography). Until the late 1990s, the prototypical language learner was a Castilian L1 speaker who had a substantial capacity to understand Catalan. Within a few years, though, the group of people taking these language courses not only grew substantially, but also diversified enormously. Coping with these transformations required a strenuous effort.

The current situation of adult language teaching and learning is quite different depending on the territory. According to the *Llocs per aprendre català* website,⁸ by 2017 there were 331 private and public institutions teaching Catalan to adults in Catalonia. The balance in the other territories was more modest. The Valencian administration offered adult language courses within the region's framework of adult education centres,⁹ and courses could also be taken through private associations and academies. In the Balearic Islands, the demand was more or less met by 14 language teaching centres belonging to the autonomous administration.¹⁰ According to the *Llocs* website, in Andorra there were five language teaching centres, only one in Northern Catalonia, and three in la Franja. Nothing was said about l'Alguer, although courses in Catalan are periodically offered by the city council, cultural organisations,

5 <http://mon.uvic.cat/siec2014/2013/05/21/hola-mon/> (last accessed: 15.03.2017).

6 CPNL website <http://www.cpnl.cat/>. History of CPNL: <http://25anys.cpnl.cat/> (last accessed: 15.03.2017).

7 <http://25anys.cpnl.cat/img/infografia25anys.pdf> (last accessed: 15.03.2017).

8 <http://aplicacions.llengua.gencat.cat/llc/AppJava/llocs.html> (last accessed: 15.03.2017). Given that this page was dependent on the Catalanian administration, it is reasonable to believe that it did not have all data available and therefore underestimated the possibilities of learning Catalan outside Catalonia.

9 <http://www.ceice.gva.es/ca/web/dgplgm/cursos-de-valenciano> (last accessed: 15.03.2017).

10 <https://www.caib.es/sites/llenguacatalana/ca/actualitat/?campa=yes> (last accessed: 15.03.2017).

the Sardinian regional authorities, etc.¹¹ Language diplomas obtained in one territory tend to be recognized in the others, although it was not until the recent 'Palma Declaration' that the Valencian authorities – which had been in the hands of a conservative party that was hostile to Catalan until 2016 – wholeheartedly committed themselves to this mutual recognition.

There is a not inconsiderable tradition of teaching Catalan as a foreign language in a number of countries. In the past, this activity derived from the presence of linguists with an interest in Catalan language in culture, often Catalans themselves, working within Romance or Hispanic departments, or the presence of Catalan migrants or exiles, especially after the Spanish Civil War. The recovery of Catalanian autonomy opened the door to more systematic and sustained support of these efforts, and by the early 1990s, the situation had improved so significantly that a handbook of Catalan studies was already deemed necessary (Bover i Font 1993, 10). The influence of the language abroad crossed a symbolic threshold in 2002 with the creation of the Institut Ramon Llull (IRL),¹² the official institution supported by several autonomous and local authorities from around the Catalan language area and specifically charged with promoting Catalan language and culture. In 2014–2015, for instance, Catalan was taught at 88 universities in 26 different foreign countries (Institut Ramon Llull 2016, 42).

4 Bibliography

- ACN (2016), *Les escoles immersives de la Catalunya Nord estan plenes*, El Punt Avui, 1.9.2016, <http://www.elpuntavui.cat/societat/article/5-societat/1000328-les-escoles-immersives-de-la-catalunya-nord-estan-plenes.html> (last accessed: 29.11.2018).
- ACN (2018), *Les famílies de l'Alguer podran escolaritzar els fills en català*, VilaWeb.cat, 13.7.2018, <https://www.vilaweb.cat/noticies/les-families-de-lalguer-podran-escolaritzar-els-seus-fills-en-catala/> (last accessed: 29.11.2018).
- Allioli (2016), *Informe sobre l'ensenyament en valencià* 263, 4–11, https://issuu.com/allioli-stepv/docs/allioli_263_juny_2016_web (last accessed: 29.11.2018).
- Anguera, Pere (1997), *El català al segle XIX. De llengua del poble a llengua nacional*, Barcelona, Empúries.
- Aparici, Artur, et al. (2011), *Els usos lingüístics a les universitats públiques valencianes*, València, Acadèmia Valenciana de la Llengua.
- Arnau, Joaquim/Vila, F. Xavier (2013), *Language-in-Education Policies in the Catalan Language Area*, in: Joaquim Arnau (ed.), *Reviving Catalan at School Challenges and Instructional Approaches*, Bristol/Buffalo/Toronto, Multilingual Matters, 1–28.
- Bover i Font, August (1993), *Manual de catalanística*, Barcelona, Publicacions de l'Abadia de Montserrat.

¹¹ <http://www.comune.alghero.ss.it/it/comunicazione/notizie/notizia/Corsi-di-catalano-di-Alghero-per-gli-insegnanti/>, <http://www.alguer.it/video/v?id=45550>, <http://www.regione.sardegna.it/j/v/13?s=320829&v=2&c=3&t=1> (last accessed: 15.03.2017).

¹² <http://www.llull.cat/catala/home/index.cfm> (last accessed: 15.03.2017).

- Canal Santos, Imma/Marco Escabosa, Anna (2016), *La immersió lingüística al sistema educatiu de Catalunya o com aprendre una llengua alhora que s'aprenen continguts*, in: Mònica Pereña (ed.), *Ensenyar i aprendre llengües en un model educatiu compartit. Metodologies i estratègies per al desenvolupament de projectes educatius i per a la pràctica docent*, Barcelona, Horsori, 79–98.
- Departament de Cultura (2016), *Informe de Política Lingüística 2015*, Barcelona, Departament de Cultura, Generalitat de Catalunya, <http://llengua.gencat.cat/web/.content/documents/informepl/arxiu/IPL-2015.pdf> (last accessed: 29.11.2018).
- DSDEN 66 (2018) = Direction des services départementaux de l'éducation nationale. Pyrénées Orientales 66 (January 2018), *Le département en chiffres. Les chiffres de la population scolarisée en 2017–2018*, <http://www.ac-montpellier.fr/dsden66/cid90140/le-departement-en-chiffres.html> (last accessed: 15.01.2019).
- Ferrer i Gironès, Francesc (1985), *La persecució política de la llengua catalana*, Barcelona, Edicions 62.
- Gabinet de Comunicació (2017), *Un 54 % d'escoles han triat el nivell avançat, un 30 % el nivell intermedi i un 16 % el nivell bàsic de plurilingüisme*, 16.3.2017, http://www.ceice.gva.es/portal-cons-portlet/htdocs/area_de_prensa/versionImprimiblePrensa.jsp?id_nota=701260&idioma=VA (last accessed: 29.11.2018).
- Gabinet de Didàctica (1997), *Índex per temes, autors i títols de tots els articles apareguts en la revista COM Ensenyar Català als adults*, *Com ensenyar català als adults* 30, 3–47.
- González, Mònica (2013), *Històrico no al TIL*, *Última Hora*, 29.9.2013, <http://ultimahora.es/noticias/local/2013/09/29/109306/marea-verde-contra-til-apoyo-huelga-indefinida-docentes.html> (last accessed: 29.11.2018).
- Institut Ramon Llull (2016), *Memòria 2015*, Barcelona, Institut Ramon Llull, http://www.llull.cat/rec_doc_mem/memoria_2015_IRL.pdf (last accessed: 29.11.2018).
- La Bressola (2007), *La realitat d'un somni: trenta anys d'escoles catalanes a la Catalunya del Nord*, Barcelona, Edicions de 1984.
- Loi n° 51–46 du 11 janvier 1951 relative à l'enseignement des langues et dialectes locaux, JORF, 13.01.1951, 483, https://www.legifrance.gouv.fr/jo_pdf.do?numJO=0&dateJO=19510113&numTexte=&pageDebut=00483&pageFin (last accessed: 15/01/2019).
- Marfany, Joan-Lluís (2001), *La llengua maltractada: el castellà i el català a Catalunya del segle XVI al segle XIX*, Barcelona, Empúries.
- Mayans, Pere/Tort, Teresa (2011), 2.3. *Ensenyament*, in: Xarxa Cruscat (ed.), *IX Informe sobre la situació de la llengua catalana (2010)*, Barcelona, Observatori de la Llengua Catalana, 60–78, <http://130.206.88.7/cruscat/web/informe/informe2010.pdf> (last accessed: 29.11.2018).
- Mayans, Pere/Tort, Teresa/Areny, Maria (2014) *Ensenyament*, in: Xarxa Cruscat (ed.), *VII Informe sobre la situació de la llengua catalana (2013)*, Barcelona, Observatori de la Llengua Catalana, 96–122, <http://www.demolinguistica.cat/arxiu/web/informe/informe2013.pdf> (last accessed: 21.01.2019).
- Mayans, Pere/Tort, Teresa/Areny, Maria (2016), 4. *Ensenyament*, in: Xarxa Cruscat (ed.), *IX Informe sobre la situació de la llengua catalana (2015)*, Barcelona, Òmnium Cultural/Plataforma per la Llengua, 96–129, <http://blogs.iec.cat/cruscat/wp-content/uploads/sites/15/2016/11/informe2015.pdf> (last accessed: 29.11.2018).
- Ministerio de Educación, Cultura y Deporte. Subdirección General de Estadística y Estudios (2014), *Los modelos lingüísticos*, in: *Las cifras de la educación en España. Curso 2011–2012*, <https://www.mecd.gob.es/dms/mecd/servicios-al-ciudadano-mecd/estadisticas/educacion/indicadores-publicaciones-sintesis/cifras-educacion-espana/2014/E1p.pdf> (last accessed: 04.03.2014).
- Monés, Jordi (1984), *La llengua a l'escola (1714–1939)*, Barcelona, Barcanova.
- Montoya Abad, Brauli (2012), *Els primers 80 anys de repressió lingüística institucionalitzada al Regne de València (1707–1787)*, *eHumanista/IVITRA* 2, 1–20.

- Nadal Badal, Oriol (1993), *Cooficialitat i bilingüisme a la Universitat Autònoma de Barcelona (1933–1939)*, Revista de Llengua i Dret 1, 129–169.
- Oliva, Lúcia/Vinent, Àngela (1980), *Català a la universitat. Entre la defensa i la passivitat*, L'hora de Catalunya, febrer, 6–13.
- Pons i Parera, Eva (2015), *The Position of Catalan in Higher Education*, in: F. Xavier Vila/Vanessa Bretxa (edd.), *Language Policy in Higher Education: The Case of Medium-Sized Languages*, Bristol, Multilingual Matters, 153–180.
- Prats, Modest (1995), *La repressió lingüística i cultural*, in: Borja de Riquer i Permanyer (ed.), *Història política, societat i cultura dels Països Catalans*, vol. 5: *Desfeta política i embranzida econòmica. Segle XVIII*, Barcelona, Enciclopèdia Catalana, 324–325.
- Pueyo, Miquel (1996), *Tres escoles per als catalans: minorització lingüística i implantació escolar a Itàlia, França i Espanya*, Lleida, Pagès.
- Renyé, Aleix (2017), *Més d'un 70 % de les famílies de Catalunya Nord volen ensenyament en català*, El Punt Avui, 2.3.2017, <http://www.elpuntavui.cat/societat/article/16-educacio/1083179-mes-d-un-70-de-les-families-de-catalunya-nord-volen-ensenyament-en-catala.html> (last accessed: 29.11.2018).
- Subdirecció General de Llengua i Plurilingüisme (2018), *El model lingüístic del sistema educatiu de Catalunya. L'aprenentatge i l'ús de les llengües en un context educatiu multilingüe i multicultural*, Barcelona, Generalitat de Catalunya, Departament d'Ensenyament, <http://ensenyament.gencat.cat/web/.content/home/departament/publicacions/monografies/model-linguistic/model-linguistic-Catalunya-CAT.pdf> (last accessed: 29.11.2018).
- Torres Sans, Xavier (2011), *Llegir, escriure i escoltar a la Barcelona del Sis-cents*, in: Joan Santanach i Suñol et al. (edd.), *Llengua i literatura. Barcelona 1700*, Barcelona, Ajuntament de Barcelona, 58–101.
- Vila, F. Xavier (2008), *Language-in-Education Policies in the Catalan Language Area: Models, Results and Challenges*, AILA Review 26/21, 31–48.
- Vila, F. Xavier (2011), *Language-in-Education Policy*, in: Miquel Strubell/Emili Boix-Fuster (edd.), *Democratic Policies for Language Revitalization: The Case of Catalan*, London, Palgrave Macmillan, 119–149.
- Vila, F. Xavier (2014), *Promoting Language Use in the School Environment: Some Lessons from the Catalan-Speaking Regions*, in: Belen Uranga (ed.), *Talking Pupils: The Arrue Project 2011. Research Results and Contributions of Experts*, Vitoria-Gasteiz, Departamento de Educación, Política Lingüística y Cultura del Gobierno Vasco, 137–154, http://www.ogasun.ejgv.euskadi.net/r51-catpub/es/k75aWebPublicacionesWar/k75aObtenerPublicacionDigitalServlet?R01HNOPortal=true&N_LIBR=051535&N_EDIC=0001&C_IDIOM=en&FORMATO=.pdf (last accessed: 29.11.2018).
- Vila, F. Xavier (2018), *Linguistic models in higher education in Catalonia: Origenes, rationales, achievements and challenges = Modèles linguistiques en enseignement supérieur en Catalogne: origines, justifications, réussites et défis*, in: Hélène Knoerr/Alysse Weinberg/Catherine Elena Buchanan (edd.), *Current Issues in University Immersion = Enjeux actuels de l'immersion universitaire*, Ottawa, Groupe de recherche en immersion au niveau universitaire (GRINU), 55–88.
- Vila, F. Xavier/Vial i Rius, Santi (2005), *Language Practices in Bilingual Schools: Some Observed, Quantitative Data*, in: Xoan-Paulo Rodríguez-Yáñez/Anxo M. Lorenzo Suárez/Fernando Ramallo (edd.), *Bilingualism and Education: From Family to School*, Munich, Lincom 263–273.
- Vila i Mendiburu, Ignasi (2012), *Balanç de la política lingüística a l'educació escolar dels governs de Catalunya (2004–2011)*, Societat Catalana 2011, 295–323.
- Woolard, Kathryn A. (1997), *Between Friends: Gender, Peer Group Structure, and Bilingualism in Urban Catalonia*, Language in Society 26, 533–560.

Oriol Camps Giralt and Aina Labèrnia Romagosa

21 Catalan in the Mass Media: The Rise of Stylebooks

Abstract: When Catalan gained access to audio-visual mass media in the 1980s/after the end of the Franco dictatorship, linguists working in the media were called on to play an advisory role and to provide explanations for their recommendations. This led to the emergence of stylebooks, addressing the need to adapt language to new communicative requirements that prescriptive guidelines did not yet foresee, which added to an atmosphere of tension between prescriptive guidelines and stylebooks. Stylebooks cover standard language, the appropriateness of language for communicative situations, registers, and the best way of connecting with an audience. Stylebooks provide their respective media with style recommendations, but with awareness that they also exert an influence on the standard language as a whole.

Keywords: language advisors, stylebooks, print media, broadcasting media, standard language, prescriptive guidelines, language models, appropriateness, register

1 From proof-readers to language advisors

In 1983, the creation of Catalunya Ràdio (Catalonia's public radio network) and TV3 (the primary television channel of Catalan public broadcaster Televisió de Catalunya), both broadcasting exclusively in Catalan, also saw the launch of their own linguistic advising units. In the print media (such as publishing houses and the daily newspaper *Avui*), linguists were employed as proof-readers, but when broadcast media appeared they became "language advisors". Though this change of name may seem like little more than a euphemism, there are substantial differences between proof-readers and language advisors. The main object of linguistic advising is not texts but the people that write them and read them, as well as other professionals that spontaneously speak in front of the cameras and microphones. The starting point of linguistic advising is therefore the language skills of communication professionals (which depend on the language spoken in the family home, their education both at school and at university, and the level of culture each one of them has acquired), and its aim is to help them generate a high quality communicative product.

An important part of the language advisor's job still consists of proof-reading. However, this task can never be separated from providing explanations to editors, giving advice on pronunciation and consulting on which word should be used to name a new concept, or on how to translate a foreign term or expression. This involves debate between linguists and journalists and other communicators. Moreover, when

<https://doi.org/10.1515/9783110450408-029>

reading on the radio or television, and even more so when speaking without a script, it is people who play the most crucial role, with their pronunciation, their way of interpreting punctuation, their intonation of sentences, or how they structure their speeches, etc. As a result, people need to be advised on their language, for the language used by broadcasters is not just theirs; it represents the language of the media they work for. This is why linguistic advising is a negotiation rather than an imposition of particular solutions.

2 Linguistic advising and stylebooks

Amid the debate on the Catalan standardisation process in the media, which has been described as a ‘national issue’¹ (Gifreu 2014, 168) since the mid-1980s, language advisors need to define the style of the media they work for, and to do so they have several resources at their disposal: the official grammar up until recently (Fabra 1933);² the body of published studies on linguistic aspects; a number of conference proceedings, such as the 2nd Symposium on the Study of Prescriptive Language in 1985; the *Proposta per a un estàndard oral de la llengua catalana* (Institut d’Estudis Catalans; hereafter IEC 1990; 1992); and several dictionaries, such as that of Fabra (1932, and successive editions), that of Enciclopèdia Catalana (1982, and successive editions), the IEC dictionary (DIEC 1995; 2007, and its online updates), etc. Given this wide range of potential sources, language advisors generally gather their observations in stylebooks, so that they can access them in the course of their work without having to start from scratch when faced with a problem. Such stylebooks also serve as a practical guide for editors and communicators, the real targets of linguistic advising and the group that contributes most to language in the media, in terms of not only content but form as well.

In this chapter, we analyse some twenty stylebooks published between 1986 and 2013, written by linguistic units in the media or intended for the media. We focus mainly on the attitudes that stylebooks adopt towards prescriptive guidelines, the development of a “standard” Catalan language, appropriateness to communicative situations, register, and how the media address themselves to their audience. For the sake of brevity, our attention here is focused on the most generic statements regarding these aspects, and therefore a detailed analysis of the specific recommendations of stylebooks is not provided in this chapter.

1 All quotes were translated by the authors of this chapter.

2 This grammar, first published in 1918, was slightly modified by Fabra in later editions – the seventh and last edition appeared in 1933.

3 Stylebooks and prescriptive guidelines: tensions

Language advisors in the broadcast media sometimes have problems with current prescriptive guidelines in Catalan. There is generally a feeling of distance between the official grammar and dictionaries, on the one hand, and “real” usage, on the other – the latter being the usage that broadcasters on the radio and newsreaders on television tend to opt for, in terms of both vocabulary and syntactic structures.

Toni Mollà (1990), writing from the point of view of Radiotelevisió Valenciana, explicitly indicates that the prescriptive code to be regarded as *correct* is based on Fabra’s work and, ‘in the Valencian Country, the Castelló subnorm of 1932’ or *Normes de Castelló* (14.1 The Language Reform, the Institut d’Estudis Catalans and the Work of Pompeu Fabra), but states that in the prescriptive guidelines there are “grey areas” that need revising. Equally, we should remember that these norms ‘are orthographical and the broadcast media need, logically, orthophonic models’ (Mollà 1990, 36–37). In other words, the available prescriptive guidelines are designed for written language, while what the broadcast media need are guidelines for spoken language.

Most stylebooks demonstrate a critical attitude towards IEC guidelines in two senses: firstly, these guidelines need to be complemented by other studies that expand on current knowledge of the language and, secondly, they need to be adapted to each medium and region. Thus, prescriptive guidelines can serve as ‘a point of reference’, ‘but, next to this ultimate authority, there are a large number of lexicographical works and studies on the Catalan language that expand, complement or further explain IEC guidelines, which cannot be underestimated when shaping standard Catalan’ (CR 1986a, 3). Even though Fabra’s dictionary (1932) ‘is the reference point, compulsory in any consultation, we also need to keep in mind that not everything in a dictionary is appropriate for a newspaper’ (DdB 1987, 26). This feeling of discomfort with prescriptive guidelines is also expressed in other print media stylebooks. For example, the stylebook of *El 9 Nou* (9nou 1991, 43) agrees, almost point by point, with the above quotation from CR. In this sense, an interesting work is the *Manual d’estil* (Mestres et al. 1995), which is meant to be a general reference work for written texts. In its introduction, the authors declare their ‘critical will to follow prescriptive guidelines’ (agreeing with most media language advisors) and refer to the ‘controversial linguistic issues or those lacking definition in the current prescriptive guidelines’ and to ‘some grammatical prescriptions that are still troublesome’.

The TV3 stylebook (1995, 6–7) only mentions the *Proposta per a un estàndard oral* (IEC 1990; 1992) and the ‘reference works – grammars and dictionaries –’ in its introduction, but the body of the work strictly consists of a phonetical, grammatical and lexical compendium. Later on, other stylebooks, such as COMRàdio (2009), applied the same formula. A critical reference to prescriptive guidelines is maintained by *Avui* (1997, 12), which states that ‘in order to correctly use language we must take prescriptive guidelines as a starting point. However, different types of media can listen to the authorised voices that have been expanding our vocabulary and syntactic

knowledge of Catalan for years', and *El Periódico* (2002, 9), which claims to be "respectful towards IEC guidelines" and explicitly indicates which of the terms and expressions it contains do not follow these guidelines. Similarly, the linguistic advising website of the Catalan Audiovisual Media Corporation (CCMA), *ésAdir*, indicates whether its usages appear in the 2nd edition of the DIEC (2007) or not, and includes the proposals of Termcat. By doing so, *ésAdir* refers to prescriptive guidelines, even acknowledging disagreements. The CCMA stylebook (2013, 334), states that its language is based on 'current prescriptive guidelines, represented by the documents approved by the IEC Philological Section and by the proposals made by Termcat and other works and sources considered a frame of reference'.

According to *El Punt* (Drou 2000, 12–13), each medium's linguistic style should take shape 'within the range of possibilities offered by prescriptive guidelines'. The Andorra TV stylebook (AndTV 2005, 6) states that language on television should approximate 'everyday language and language of general usage', 'always taking Catalan prescriptive guidelines as the starting point'. IB3 (2006, 15) stresses that Balearic forms 'accepted by prescriptive guidelines' have been accepted in its stylebook. UIB (2009, 17) states: 'In composing this manual, we have based ourselves primarily on the proposal of the oral standard' made by the IEC 'and applied it to the Balearic Islands'. The stylebook of the Catalan News Agency (ACN 2010, 150) establishes that 'each professional shall respect prescriptive guidelines according to their linguistic variant'.

Worthy of particular mention is the Vilaweb stylebook (VW 2006, 5), which declares that its pages, 'following both Fabra's general linguistic guidance and the reasoned proposals that are inspired by it and give it continuity, do not feel the need to question the good work that has already been done'. For this reason, they 'do not intend to address the language model nor the language grammar, but to offer stylistic criteria'. VW, then, not only acknowledges no conflict with current prescriptive guidelines, but actually supports their main source, which they choose to develop in their own way. Similarly, the Acadèmia Valenciana de la Llengua or Valencian Academy of Language's media stylebook (AVL 2011, 6) quotes as prescriptive sources its own grammar (AVL 2006a) and dictionary (AVL 2006b). In this work there is no mention of any conflict either, due to the fact that it intends to disseminate its own prescriptive guidelines.

4 Relieving the tensions: concepts

Going back to the 1980s and 1990s, language advisors have resorted to a number of concepts in the attempt to find a solution to conflicts with prescriptive guidelines, including standard language, appropriateness, registers, and the need to take the intended audience into account in order to achieve good communication.

4.1 The standard language

The concept of standard language had been a subject of study at universities for years, but it spread beyond academic circles due to the discussion of language in the media. Indeed, the standard variety of language has always been connected with the purposes of education (school) and information (the media). In the *Orientacions per a l'ús de la llengua a Catalunya Ràdio i RAC*, the standard variety is described as follows:

'The radio medium requires the use of a language common to most speakers, without unwarranted loan words, reduced domain dialecticisms, and archaisms improper for oral style. This common language, which is called *standard*, is the most neutral and usual version of the language, and the one that the speech community as a whole can identify most easily.' (CR 1986a, 1)

The stylebook of the *Diari de Barcelona* also defines the standard as 'the variety common to most speakers of a language', which 'comprises everything that is general [...] and at the same time rejects any local peculiarity' (DdB 1987, 25). However, the stylebook is also quick to clarify that this common language was compiled taking Central Catalan as its basis due to Barcelona's demographic, economic and cultural importance over many centuries. In these Barcelona-based media, the 'standard construction' is built up as if it were another register within the Central dialect. The standard language intended for the masses is made up of everyday language, dispensing with archaisms, dialecticisms and 'unwarranted loan words', but accepting some 'historical *Castilianisms*' and other forms of general usage that do not appear in Fabra's dictionary (1932).

Toni Mollà (1990, 37), for his part, tries to fit Valencian into the framework of the general standard outline in prescriptive guidelines, and defends "the use of a regional linguistic variety that corresponds as much as possible with the media's area of coverage":

'As a consequence, we will have to banish from the linguistic model(s) of these media not only linguistic structures not belonging to the prescriptive codification (such as **tindre que* ['to have to', instead of *haver de*]), but also those forms that, even though they are comprised within the diasystem and are accepted by the prescriptive guidelines, are typical of geographical areas outside the scope of that specific media. Therefore, in the case of Valencian media, it will be necessary to banish those linguistic forms characteristic of other areas of the Catalan Countries' (Mollà 1990, 37).

In this context of conceptual struggle, the first two booklets of the *Proposta per a un estàndard oral de la llengua catalana. I Fonètica* (IEC 1990) and *II Morfologia* (IEC 1992) were published. They were meant to provide some guidelines on specific issues to complement the official grammar (Fabra 1933) by means of instructions on phonetics and morphology that took the different geographical varieties of Catalan into account, and with a compositional and polymorphic approach to the standard language. Nevertheless, this approach was still far from the one that linguists in the

media would have hoped for, mostly because, instead of establishing some specific models that would be easy to follow – important for the urgency of the media –, they outlined the conceptual framework in which a standard oral language in Catalan could be developed.

All subsequent stylebooks claim to use the standard language; those in the print media, without further comment, except for the stylebook of VilaWeb, which is in favour of ‘building a national standard language, not a regional one, suitable for the whole linguistic domain’ (2006, 5). For their part, the broadcast media specify on which variety they base their concept of standard language: TV3 (1995) mainly uses the Central variety, without excluding others; AndTV (2005), Northwestern and Central; IB3 (2006) and UIB (2009), Balearic; BTV (2008), COMRàdio (2009), ACN (2010) and CCMA (2013) accept “all varieties”. This means that over two decades there has been an evolution in how the standard language is conceived, which has come closer to the compositional and polymorphic view of the IEC, even though each of these media mainly seeks to construct the part of the standard that corresponds to its area of coverage. On this point, the exception is the media stylebook of the AVL (2011, 10), which is skeptical towards the idea of the standard variety: ‘we do not speak of the existence or not of this variety. We simply describe the usages of Valencian in public audio-visual communication’. Nevertheless, this stylebook contains many references to the standard, though it can be inferred that it only refers to the regional Valencian standard.

4.2 Appropriateness

Next to the concept of standard, there is the concept of appropriateness, which very often appears in relation to correctness:

‘We need to relativise the use of the word *correctness*. It would be convenient to delimit the criterion of correctness of a specific variety according to its context of realisation. It is in this sense that perhaps the word *correctness* should be replaced by *appropriateness*. As a result, it can be said that a given linguistic form will be *correct* (appropriate) when it is *fitting* in that communicative situation in particular’ (Mollà 1990, 17).

While correctness refers to prescriptive guidelines initially conceived for written language, appropriateness is a concept that is still developing and that points to different aims. These aims relate not only to the variety, as we have just seen, but also to the media:

‘The main aim of these guidelines is to achieve an appropriate language for the media, one that is clear, direct, simple, and genuine, with preference for more natural and spontaneous forms of spoken language [...] to establish a framework in which linguistic expression is appropriate, so that each word occupies its place’ (CR 1986a, 1).

Appropriateness is also related to the intended audience, according to AndTV (2005, 6), IB3 (2006, 21), and AVL (2011, 7) (cf. § 4.4.).

Later, the concept of appropriateness started to be introduced into prescriptive guidelines as well. The notion is referred to twice as a specialised term, albeit along with other usages, in the introduction to the second fascicle of the *Proposta per a un estàndard oral de la llengua catalana. II, Morfologia* (1992). In the first of these occurrences, there is the idea – which first starts to appear in stylebooks – of linguistic usages appropriate to specific tasks in the media, whereas, in the second, appropriateness of linguistic forms refers to the communicative situation:

‘With the proposals for an oral standard we recover and dignify a series of usages appropriate to tasks that had been interrupted for years or that simply did not exist previously’ (IEC 1992, 10).

‘The first label (*it is proper* [in Catalan, *pròpia*]) means that the form in question is the one that is fully appropriate and can be used without limitations in all registers; the second one (*it is admissible*) presupposes the existence of another form – the one that is *proper* and common –, which is always preferable, and implies that the *admissible* form is merely tolerated and that it is advisable to replace it with the *proper* form in the most typical situations of the standard language’ (IEC 1992, 11).

The UIB and the CCMA also relate appropriateness to the communicative situation:

‘Language correctness is not enough; language has to be appropriate to the situation as well. This means that a speech, a text, may be perfectly correct but inappropriate if it does not match the context, the communicative situation (formal/informal) in which it is used. Therefore, it is the communicative situation that sets out the rules for the sort of language we should use’ (UIB 2009, 18).

‘We adjust our expression to the topic, to the potential audience, to the will for proximity between the speaker and the interlocutor or the audience, to whether there is a direct interaction with those who receive the message, to whether the intervention is spontaneous or planned, or to whether it is written or oral language, among other factors’ (CCMA 2013, 340).

This evolution is highly important, to the extent that the concept of appropriateness to the communicative situation has become the paradigm of the new Catalan official grammar, the *Gramàtica de la llengua catalana* (GIEC 2016).

4.3 Registers

The idea of appropriateness to the communicative situation leads us naturally to the concept of register. The *Orientacions* of Catalunya Ràdio explain register as follows:

‘Talking about medicine is not the same as talking about football; each topic requires differentiated terminology and expressions. Writing an article on a basketball match for a newspaper is not the same as commentating on it live on the radio; the difference between written and spoken Catalan is pretty clear. Writing an advertising text is not the same as writing the news for a news

programme; the purpose of communication involves specific appropriateness for each case. And the degree of formality in a documentary programme is not the same as that in a magazine programme; the first does not allow colloquialisms, whereas in the second no stylistic stiffness is allowed, even if it seems genuine. These different situations in which a language can be used form registers, which can be multiple, as we have seen, according to the topic, the written or oral mode of the media, the communication purpose, and the degree of formality' (CR 1986a, 2).

The UIB (2009) and the AVL (2011) refer to registers according to the type of programme:

'Thus, it is important to distinguish between performing a formal and non-spontaneous monologue and participating in a discussion among two or more people, in which the speech is somewhat informal and spontaneous. In the first case we use the standard variety, whereas in the second case certain usages of informal speech are admissible' (UIB 2009, 19).

'A classification of the typology of the programmes in accordance with the interpersonal tone needs to distinguish between monologues and dialogues. In the first case, there is no interaction between sender and receiver; rather, communication is unidirectional. In these cases, the register is usually formal, as in news programmes or reports. In the second case, when a sender and a receiver interact, the variety tends to be less formal (debates, talk shows, informal interviews, entertainment programmes ...)' (AVL 2011, 11).

4.4 Taking the recipients into account

Communication between the media and the audience mainly takes place by means of language. This is why the media should use a language close to their target audience, in terms of both variety and register.

'The only way for speakers of the whole linguistic domain to feel recognised and accepted in the media we were talking about would be if the linguistic vehicular model(s) does not drift too far from the colloquial-dialectal languages that these speakers know and use' (Mollà 1990, 17).

'If the term that is usually used is already Catalan, replacing it does not confer any more Catalan-ness on the sentence; on the contrary, it reduces its naturalness and distances it from the characteristic language of the radio medium. This way, it is hard for listeners to recognise their own language and they will find it not very understandable and difficult' (CR 1986a, 5).

That being said, recognition is mutual: the medium recognises its audience by using the variety that is closest to the audience – 'the linguistic solutions from the broadcasting sphere of the media have to be particularly highly valued' (COMRàdio 2009, 19) – and the audience recognises the language of the medium, which is appropriate 'if its viewers identify with it and feel it as their own' (AndTV 2005, 6); and 'if viewers or listeners recognise it as their own and it agrees with the norms of traditional usage' (IB3 2006, 21). Moreover, in order for communication to work, the message must also be 'understandable' and 'close to everyday language and the language of general usage' (AndTV 2005, 6), using 'common words' (VW 2006). The print media merely

demand a language that is ‘attainable for an average reader’ (DdB 1987, 25), whereas the broadcast media opt for ‘direct sentences that assist the understanding of the text’ (BTV 2008, 82), and insist that the ‘listeners cannot wind back, so they have to be able to understand and retain information as soon as they hear it’ (CR 1986a, 5); ‘the viewer or listener should grasp the message as soon as they hear it’ (UIB 2009, 22); ‘the message of oral communication is not meant to be repeated, but to be understood as soon as it is broadcast’ (AVL 2011, 7); ‘among different possible correct forms, we choose the most understandable for everyone’ (CCMA 2013, 336).

5 Media stylebooks and the standard language

What, then, are media stylebooks? According to *El Punt* (Drou 2000, 12), ‘Very often, stylebooks have been regarded as practice grammars that dissent from prescriptive guidelines. They do not have to, but they cannot be restricted to glosses of prescriptive texts either’. The notion that stylebooks are ‘neither dictionaries nor grammars’ is repeated in several of these works (9nou 1991, 17; *El Periódico* 2002, 7; CCMA, *ésAdir*), indicating that the model they suggest ‘is not a model for the whole language’ (9nou 1991, 17), but one mainly addressing journalistic language, which ‘is a specialised language’ (Avui 1997, 13), ‘a specific textual typology’ (9nou 1991, *ibid.*). Thus, from the standpoint of both those that refer to a language model and those that refuse to even consider such a thing, stylebooks ‘offer stylistic criteria’ (VW 2006, 5), addressed to the media, without denying that they exert an indirect influence on the standard language as a whole.

6 Bibliography

- 9nou = Coromina, Eusebi (1991), *El 9 Nou. Manual de redacció i estil*, Vic, Eumo/Prensa d'Osona.
- ACN = Ferré Pavia, Carme/Nogué Regàs, Anna (2010), *Llibre d'estil. Agència Catalana de Notícies*, Barcelona, UOC.
- AndTV = Puigdomènech i Farell, Laura/Solís Obiols, Marina (2005), *Proposta de llibre d'estil per a Andorra Televisió*, Andorra la Vella, Govern d'Andorra, Servei de Política Lingüística.
- AVL = Acadèmia Valenciana de la Llengua (2006a), *Gramàtica normativa valenciana*, València, Publicacions de l'Acadèmia Valenciana de la Llengua.
- AVL = Acadèmia Valenciana de la Llengua (2006b), *Diccionari ortogràfic i de pronunciació del valencià*, València, Publicacions de l'Acadèmia Valenciana de la Llengua.
- AVL = Castillo Ventura, Carmen, et al. (2011), *Llibre d'estil per als mitjans audiovisuals en valencià*, València, Acadèmia Valenciana de la Llengua.
- Avui (1997), *Llibre d'estil del diari Avui*, Barcelona, Empúries.
- BTV = Ferré Pavia, Carme (2008), *Barcelona tv. Llibre d'estil*, Barcelona, UOC.
- CCMA = Corporació Catalana de Mitjans Audiovisuals (2013), *Llibre d'estil de la Corporació Catalana de Mitjans Audiovisuals. Guia editorial. Manual d'ús*, Barcelona, Generalitat de Catalunya,

- Corporació Catalana de Mitjans Audiovisuals, <http://www.ccma.cat/llibredestil/> (last accessed 27.02.2017).
- CCMA, *ésAdir*, <http://esadir.cat> (last accessed 27.02.2017).
- COMRàdio = Ortega, Rudolf/Brunat, Gemma (2009), *Llibre d'estil de COMRàdio: manual de llengua de les emissores municipals*, Barcelona, Diputació de Barcelona.
- CR = Catalunya Ràdio (1986a), *Orientacions per a l'ús de la llengua a Catalunya Ràdio i RAC*, Barcelona (unpublished).
- CR = Catalunya Ràdio (1986b), *Principals faltes davant el micròfon*, Barcelona (unpublished).
- CR = Catalunya Ràdio (1992–2001), *Orientacions lingüístiques*, Barcelona (unpublished) (adaptation of the two previous works).
- DdB = Diari de Barcelona (1987), *Un model de llengua pels mitjans de comunicació. Llibre d'estil del Diari de Barcelona*, Barcelona, Empúries.
- DIEC = Institut d'Estudis Catalans (1995), *Diccionari de la llengua catalana*, Barcelona/Palma de Mallorca/València, Enciclopèdia Catalana/Edicions 62 /Publicacions de l'Abadia de Montserrat/ Moll/Edicions 3 i 4.
- DIEC = Institut d'Estudis Catalans (2007), *Diccionari de la llengua catalana*, Barcelona, Edicions 62/ Enciclopèdia Catalana.
- Drou, Pere (2000), *Llibre d'estil d'El Punt: llengua 2*, s.l., Hermes Comunicacions.
- El Periódico (2002), *Llibre d'estil*, Barcelona, Ediciones Primera Plana.
- Enciclopèdia Catalana (1982), *Diccionari de la llengua catalana*, Barcelona, Enciclopèdia Catalana.
- Fabra, Pompeu (1932), *Diccionari general de la llengua catalana*, Barcelona, Llibreria Catalònia.
- Fabra, Pompeu (?1933), *Gramàtica catalana*, Barcelona, Institut d'Estudis Catalans.
- Gifreu, Josep (2014), *El català a l'espai de comunicació. El procés de normalització de la llengua als mitjans (1976–2013)*, Bellaterra/Castelló de la Plana/Barcelona/València, Servei de Publicacions de la Universitat Autònoma de Barcelona/Publicacions de la Universitat Jaume I/Universitat Pompeu Fabra/ Publicacions de la Universitat de València.
- GIEC = Institut d'Estudis Catalans (2016), *Gramàtica de la llengua catalana*, Barcelona, Institut d'Estudis Catalans.
- IB3 = Company, Catalina/Puigròs i Caldentey, Maria Antònia (2006), *Llibre d'estil d'IB3*, Palma de Mallorca, Consorci per al Foment de la Llengua Catalana i la Projectió Exterior de la Cultura de les Illes Balears.
- IEC = Institut d'Estudis Catalans (1990), *Proposta per a un estàndard oral de la llengua catalana. I Fonètica*, Barcelona, Institut d'Estudis Catalans.
- IEC = Institut d'Estudis Catalans (1992), *Proposta per a un estàndard oral de la llengua catalana. II Morfologia*, Barcelona, Institut d'Estudis Catalans.
- Mestres, Josep M., et al. (1995), *Manual d'estil. La redacció i l'edició de textos*, Vic/Barcelona, Eumo/ Universitat de Barcelona/Universitat Pompeu Fabra/Associació de Mestres Rosa Sensat.
- Mollà, Toni (1990), *La llengua dels mitjans de comunicació*, Alzira, Bromera.
- TV3 (1995), *El català a TV3: llibre d'estil*, Barcelona, Edicions 62.
- UIB = Picó, Neus/Ramon, M. Magdalena (2009), *Llibre d'estil per als mitjans de comunicació orals i escrits*, Palma, Consell Insular de Menorca/Universitat de les Illes Balears.
- VW = VilaWeb (2006), *Llibre d'estil*, s.l., Vilaweb.

22 Terminology and Neology

Abstract: As the practice of compiling terms in a field of specialisation, terminology is an essential activity for specialised communication. All specialised fields have their own terminology, and all specialists use specific terms to transfer their knowledge of the field. Because of its precise nature, terminology makes expert communication more efficient. However, using and compiling terminology can serve a number of different social purposes, from international communication to promoting the use of a language in language planning. Catalan terminology is an example of an activity serving the social promotion of a language. The availability of terminology ensures that a language can be used in all spheres of communication. Catalan terminology has been the subject of an organised project within the framework of the linguistic normalisation plan launched in 1981. It is currently an exemplary model for other communities with similar objectives. Catalonia's contribution to international terminology has not been limited to its practice, encompassing both theoretical innovation and the creation of technologies for processing terminology.

Keywords: linguistics, terminology, neology, terminography, Catalan language, Catalan terminology

1 Purpose and scope of terminology and its position within linguistics

Terminology is the field concerned with analysing and compiling specialised terms. It has a long-standing tradition as a practice that involves compiling terms by fields of knowledge and developing glossaries and dictionaries. The need for compilations of the lexical elements required to name scientific, technical and specialised concepts in general has existed for centuries. Terms ensure that knowledge is represented and transmitted precisely.

Despite the long history of the practice, terminology as a discipline focused on terms has emerged quite recently. Academic contributions strictly focused on terms began in Western Europe in the first third of the 20th century as a consequence of the interest in institutionalising a discipline that was essentially practical. We could claim that Eugen Wüster (Wieselburg, 1898–1977) was the mastermind behind the institutionalisation of terminology and its inclusion as a field of knowledge, as well as its recognition as a social need linked to the internationalisation of science and technology.

<https://doi.org/10.1515/9783110450408-030>

Activity in terminology is associated with the need to disseminate scientific and technical knowledge and the products derived from this knowledge beyond the borders of a given speech community. The rapid development of new technologies in the late 19th and early 20th centuries required defining and compiling the names of new products on a permanent basis, a need that the scientific nomenclatures and international symbols used in previous centuries couldn't meet anymore. In the hands of experts, particularly for commercial purposes, terminology became an extremely important tool in emerging economies. The need to transfer technology without any communicative ambiguity led industrialists and technology experts to start normalising specialised names. Thus began the institutionalisation of terminology *per se*. There were three main facets to this institutionalisation:

- a) As an industrial activity associated with ISO (International Standardisation Organisation): an international committee was created within ISO, the CT-37, charged with standardising the concepts and methods of terminology.
- b) As a specialised subject with an international scope: a terminology information centre (Infoterm) headquartered in Vienna was created with the sponsorship of UNESCO. Infoterm adds dictionaries, lexicons and vocabularies produced around the world to a database and disseminates information on terminological activities, which are increasingly plentiful.
- c) As a field of knowledge: terminology has been introduced into university education and a theory has been developed to explain the basis of terminology from a scientific (or purportedly scientific) standpoint, initially at the University of Vienna.

The only terminology proposal that spread during this era came from Central Europe, specifically Vienna. It was developed by Eugen Wüster, an Austrian engineer who was profoundly interested in unambiguity in the transmission of specialised knowledge in the context of economic internationalisation.

Linguistics, which at that time revolved around the emerging structuralist movement, had little to contribute to this new activity. The theoretical underpinnings of terminology were far from linguistic approaches in both their basic premises and in their consideration of how a language evolves or should evolve. Unlike the words of a natural language, terms are mere designations of supposedly universal concepts. Unlike linguistics, terminology is only interested in terminological units as labels of concepts, and intervening in the evolution of language by defining fixed names is an unavoidable part of terminological activity.

Therefore, it is obvious that linguistics, which at that time set out to describe natural languages and their natural evolution, had to distance itself from such a distinct approach to language. Linguistics today, albeit on a smaller scale, has ignored terminology and generally considered it more a practice than a field of knowledge.

Since the 1960s, terminological activity, centred on the social promotion of languages instead of internationalisation, has gradually led terminology to approaches

that are closer to linguistics, but linguistics with a sociological foundation. In these new approaches, terms are not mere designations but units from the lexicon of languages. The specificity of the terminological units lies in their semantic and pragmatic conditions; they are units with a given meaning linked to the way they are used. This has even led to a consideration of the possible *values* of lexical units (Cabr   1999; 2003).

2 Evolution of terminology

The first proposals to regulate the use of language in order to promote the status of said language appeared much earlier than the 20th century. However, as regards terminology, it was Quebec's *Charte de la langue fran  aise* (1977) that had the most profound effect on the homogeneity of terminological activity associated with the international transfer of products and knowledge. Thus, a new scenario emerged for terminology: terminological activity focused on promoting a language within society.

The working methods established by the ISO norms were not the most appropriate for this new undertaking, and the ideas and underpinnings of terminology developed by W  ster even less so.

Indeed, rather than ignoring the social origins of terms, Quebecois terminology had firm sociolinguistic underpinnings. Seeing social usage as the ultimate objective of terminological undertaking, they chose terms according to their practical usage potential rather than the "best" possible term. The goal was to update the specialised lexicon in general, including scientific and technological terminology to make the language, in this case Quebec French, appropriate for all uses, ranging from the professional to the technical and the scientific.

Therefore, we can observe two main shifts from the terminology developed in Central Europe: first, the terms had to follow the physiognomy of the language and give the appearance of genuineness whenever possible; secondly, their use should not exclude linguistic variation, a concept which had previously been fiercely argued against in favour of a single form for each language. Catalan terminology essentially had to align itself with this fundamentally communicative option.

This new perspective opened the door to concepts of terminology which until that point had been left unexplored. In a similar vein, thanks to the development of linguistics, sociolinguistics and cognitive sciences, doubt was cast on the view of terms as merely functional, ahistorical, asocial and exclusively denotative units. Terms came to be regarded as units of representation and communication common to natural languages, with connotations just like all lexical units. The reduction of terminological units to mere labels, as upheld by W  ster and his followers, had to give way to a more pragmatic approach.

Cognitive sciences favoured specialised knowledge linked to discourse. In this context, the categorisation of reality into concepts is a process that could be carried out on the margins of culture.

Generative Linguistics stressed acknowledging universal knowledge of language and the fact that this universal competency is projected in particular competencies associated with natural languages.

Terminology tends to be gradually acquired by specialists in a given domain through the process of learning a subject or activity. We could say that specialists are the natural users of terminology. Without terminology, the discourse of any given specialised field could not take place.

Real terminological practice, as it expands socially, is diversified to deal with users' specific needs. This opens a broad array of potential scenarios of terminological work which can be summarised in the following:

- Terminology at the service of international normalisation and communication.
- Terminology at the service of promoting a language.
- Terminology at the service of translation and interpretation.
- Terminology at the service of knowledge management through the organisation of documentation.
- Terminology for the dissemination of specialised knowledge.
- Terminology at the service of automatic language processing.

Today, terminology is an increasingly recognised field of study and work. University programmes, primarily translation programmes, have added general terminology courses on the principles of terminology, terminology management and terminological problem-solving in the context of translation. The importance of documentation studies has also been recognised. However, terminology has had and continues to have a scant presence in linguistics and sociolinguistics programmes, despite the fact that it is one of the ubiquitous components of all natural languages, even those in the oral tradition.

However, the practice of terminology has spread rapidly within national and international bodies, and in companies interested in quality management. Such institutions engage in a host of terminological activities, something that confirms that terminology is now a discipline in itself.

3 Applied terminology or terminography

The ultimate expression of the applied facet of terminology is the development of glossaries and dictionaries on specialised topics.

The terminographic process, which some authors categorise under the umbrella of terminology management, can be defined as an ordered series of consecutive operations based on a series of underlying principles which must be respected to ensure quality terminology. These principles include:

- Creating terminology is not translating.
- Terminology should not be mistaken with neology (although this can be used as a resource for suggesting equivalents).

- Creating terminology presupposes respect for its underlying principles (which we view as terminological unity) and the use of a methodology that suits the needs pursued by the creation of a terminological resource.
- Terminological work is always descriptive to begin with (although it can later be subject to normalisation).
- A unit is a term in a language and in an area of specialisation (its use in the discourse and its belonging to an organisation of concepts must be taken into account).
- The terminology of a glossary is not predetermined but is determined according to the project being carried out.
- All terms have a real source that justifies their use.
- All terms are associated with a grammatical category.
- All terms are associated with a definition appropriate to its use in a given field.
- All terms are associated with a pragmatic value when they are chosen for a terminographic resource.
- All terms may have denominative variations with or without cognitive consequences.
- Terms from different languages will be more or less equivalent according to the proximity of their conceptual systems.
- The harmonisation of conceptual systems falls under the category of prescriptive (not descriptive) terminology.
- A term is a unit of form, content and social function.
- The form of a term is doubly systematic (in relation to both the language and the specialised language).
- Terms are connected by (different kinds of) relations and are structured systems in each area of specialisation (systems of concepts).

Based on these principles, the most important activities in the terminographic process are the following:

- Establishing the conceptual structure of the field.
- Setting up a specialised corpus that is representative of the sphere of work.
- Identifying the terminological units in this corpus.
- Creating a terminological archive in the form of a database.
- Developing terminological records.
- Analysing the identified terminology.
- Problem-solving or reporting these problems to the relevant organisations: conceptual and denominative variation, terminological lacunae, conceptual and denominative ambiguities, lack of equivalence or of an equivalent in multilingual projects, etc.

These operations can be performed manually or using the resources and tools that have been developed by linguistic engineering, including:

- As reference resources:
 - specialised multilingual text banks,
 - banks of terminological data,
 - parallel corpora on specialised topics,
 - banks of specialised knowledge,
 - ontologies,
 - thesauri, and
 - dictionaries.
- As working tools:
 - database management systems,
 - text classifiers,
 - online text searches in a given field of specialisation,
 - term extractors,
 - systems to help develop definitions, and
 - search engines for equivalents in parallel or multilingual corpora.

4 Catalan terminology: Origins and evolution

In minority languages, and especially in languages in a minoritised position, terminology fulfils a symbolic purpose in addition to its primary purpose of representing and communicating specialised concepts. Thus, from the standpoint of social representation, the fact that a language has terminology symbolises that it can be used in all communication situations.

When the production of terminology by specialists in different fields of knowledge is irregular, especially because of the use of other languages of communication in specialised fields, neological work becomes common and legitimate in terminology, and the systematic, institutional organisation of terminology becomes absolutely essential. The availability of normalised terminology ensures that the language remains equipped with all the lexical resources it needs, and further ensures that these remain up to date.

In Catalonia, despite the historical events which affected the language throughout the 20th century, neither the production of scientific-technical discourse in Catalan nor the concern with terminology have ever ceased. Members of Catalan-speaking society have constantly been launching both individual and joint linguistic initiatives since the first quarter of the 20th century. However, the communicative function of the specialised fields has sometimes been limited to quite small circles, such as during the Franco dictatorship, when Catalan was banned as a public language in administration, education and the media. Despite this ban, there were experts who continued writing in Catalan and publishing their works underground for a small readership – for example, the members of the Institut d'Estudis Catalans (IEC, Institute for Catalan Studies), who later promoted the Lexicography of Sciences Coordination Committee.

One of the clear forerunners of the activity of setting terminology can be found in several dialectological studies published in the *Butlletí de Dialectologia Catalana* (1913–1936). However, perhaps the most important was the creation of the Lexicography of Sciences Coordination Committee in 1978, whose members included such prestigious organisations as the Academy of Medical Sciences of Catalonia and the Balearic Islands, the College of Pharmacists of Barcelona, the College of Doctors of Barcelona, the Torrens-Ibern Foundation, the Catalan Natural History Institution, the Catalan Biology Society and the Catalan Society of Physical, Chemical and Mathematical Sciences, in addition to student associations from the groups promoting Catalan at the universities in Barcelona. One of the Committee's activities was the publication of the *Fulls Lexicogràfics* (1978–1982), a collection of leaflets in which renowned experts in their respective fields debated the proper forms of variable or borrowed terms. In these debates, the experts drew from 'philologists', colleagues from the Secció Filològica ('Philology Section') of the IEC or with close ties to it, including Joan Bastardas and Jordi Carbonell. A total of eleven *Fulls Lexicogràfics* were published.

The purpose of this scientists and technology experts group, most of whom were members of the IEC, was twofold: first, to promote the use of Catalan in science and technology; and secondly, to establish a genuine international terminology. This double goal signalled a desire for consensus and precise knowledge of the rules of word formation.

Another forerunner to the institutionalisation of terminology was the appearance of the *Gran enciclopèdia catalana*, one of the works regarded as 'local', which was first published in instalments in 1968. The series was the first major update of the lexis, and thanks to the contributions of a large number of experts in different fields, it adapted the Catalan language by putting into circulation a huge number of terms corresponding to new concepts that had emerged over the years in which the Catalan language had not been officially recognised.

The institutionalisation of Catalan terminology was possible with the arrival of democracy in Spain and specifically with the recognition of Catalonia's autonomy and the creation of the Directorate General of Language Policy in 1980. The TERMCAT Terminology Centre was created in this context in 1985. However, thanks to resistance from the experts who had devoted themselves to Catalan in the previous decades, by 1985 official terminological activity already had discursive foundations to base itself on and was not limited to creating terms in Catalan *ex novo*.

Prior to the creation of TERMCAT, in 1984, the Catalan government's Directorate General of Language Policy had held a seminar to reflect on terminology, which highlighted the need to officially organise Catalan terminology for the first time. University professors, authors and publishers of dictionaries, as well as scientists and technology experts who had previously made their interest in terminology known, all participated in this seminar. They debated what the most appropriate language policy model should be for Catalan terminology, bearing in mind that Catalan extended

beyond the strict borders of the government of Catalonia. Several organisational models were considered, and finally the language normalisation model of French in Quebec was chosen, since it was the only one that had attached particular importance to terminology in the context of promoting a minoritised language.

Therefore, based on the initial idea of the Catalan government's Directorate General of Language Policy, the creation of an official entity responsible for different tasks was promoted. These tasks included developing and coordinating terminology, establishing criteria for creating terms and guidelines on creating neologisms, overseeing the normalisation of new or variable forms, and disseminating the agreed-upon terms. In order to carry out these responsibilities, the Catalan government and the IEC created TERMCAT in 1985. The participation of the public authorities ensured not only the survival of the centre but also the necessary coherence between the planning of the normalisation of the language as a whole and the terminological activity that was to serve this normalisation. The involvement of the IEC provided a linguistic endorsement.

In 1994, TERMCAT established a consortium made up of the Catalan government, the IEC and the Consortium for Language Normalisation (a normalisation body within the municipalities), and thus became its own legal entity. This change allowed the terminology centre to operate more independently and to generate its own resources through the provision of services to companies and bodies within the administration itself, as well as participation in projects. However, universities and research organisations were absent from this consortium.

Within TERMCAT, the official organisation charged with terminological normalisation in the Catalan language is the Supervisory Council, a standing body whose participants include members of the IEC and TERMCAT and experts from the fields of technology, science and humanities. The presence of members appointed by the IEC seeks to ensure that terminological normalisation is integrated into all normalisation activities in the Catalan language. The participation of the members of TERMCAT, meanwhile, ensures that valid terminological methods are used and appropriate linguistic, terminological and sociolinguistic criteria are applied. Finally, the representation of experts from different fields of specialisation seeks to guarantee that the proposals meet the specific existing needs and real usage of the terms, as well as serving as an endorsement for the future implementation of the proposed terms. Therefore, the goal is to bring together all three facets of normalisation (linguistic, terminological and usage) in the Supervisory Council in order to ensure the successful implementation of the normalised terminology.

The council debates the terms that are the subject of intervention and decides on a name for each term based on terminological policy criteria, which strive to systematise the entire process, with the goal of lowering the number of denominative possibilities as the best way to ensure unambiguous communication on academic, professional and economic matters. The approved terms are publicised via a database called Neoloteca (<http://www.termcat.cat/neoloteca>).

Terminological normalisation within the framework of Catalan linguistics, that is, as defined by TERMCAT (2006, 21), is strictly applied to specialised fields, not to the entire corpus of the language. That being said, it is applied to an extremely broad range of areas, encompassing everything from scientific and technical fields to the humanities and social sciences:

‘[...] normalisation is viewed as a process of setting the most appropriate linguistic forms for the concepts in the different fields of knowledge in a given language. Just like any other form of normalisation, this process strives to lower diversity, in this case denominative diversity, to facilitate precision and appropriateness in communication exchanges. [...] therefore, normalisation seeks to foster not only correctness but also appropriateness and conceptual precision.’

Despite all the individual, collective and institutional efforts to normalise and promote the use of Catalan terminology, the current situation is that Catalan is caught between two pressures: pressure from English in the scientific-technical fields, and pressure from Spanish in colloquial usage. These pressures manifest themselves in grammatical and lexical interference from Spanish, and primarily in lexical influence from English, especially in areas of thematically restricted use. What is more, the Catalan legal framework does not provide sanctions for violating the use of official terminology, since their use is only mandatory in the administration of Catalonia and the media which depend on it (Decree 107/1987, cf. Secretaria de Política Lingüística 2002). Indeed, it should be noted that despite the fact that TERMCAT is a referent for all the Catalan-speaking regions, it is only official within the administrative borders of Catalonia.

The current situation of Catalan terminology and the changes that have taken place since TERMCAT was created more than 30 years ago give us cause for reflection. Looking to the future, we should consider what the optimal model for organising Catalan terminology as a whole should look like. It is important to bear in mind that Catalan terminology goes far beyond the official terminology developed by TERMCAT. This terminology is supplemented by the terminological work of the IEC, which is not restricted to its participation on the Supervisory Council, and the terminological production in Catalan carried out by other organisations as well, such as university language services.

Within this framework, in 2002 the Catalan Terminology Association (ACATERM) was founded with a pluralistic membership, and in 2008 it became a subsidiary of the IEC, affiliated with the Secció Filològica, with the new name Catalan Terminology Society (SCATERM). According to its bylaws (Societat Catalana de Terminologia 2011), ‘its mission is to foster the dissemination of terminology in the Catalan language in the scientific and technical fields, to serve as a platform to promote the recognition and professionalisation of the field of terminology, and to promote the provision and exchange of information on terminological activities between terminology users and professionals’. Furthermore, every year SCATERM organises a Catalan terminology conference which aims to serve as a forum for fostering relationships among professionals working in the field.

5 Current status of Catalan terminology

Terminological activity in Catalan is truly collective. On the one hand, it is the outcome of the social-professional needs of scientists, technologists and specialists in general to represent and communicate their specific knowledge in the Catalan language. However, it is also the product of an entire society's interest in updating the lexicon of the language in order to secure its status as a language of culture at all levels. Unlike in many other countries, the institutions in Catalonia have also had the sensitivity and wisdom to channel these needs into operating organisations, such as TERMCAT and the IEC.

In fact, the creation of TERMCAT was a major step forward in the organisation of Catalan terminological activity. Although it is true that, as discussed above, terminological work existed before TERMCAT, the terminology centre instigated a systematic work process. Its efforts in terminological production (either directly or in conjunction with other organisations), promoting terminological studies and setting Catalan forms has been notable since 1985. This is supported by the centre's productivity data, which were compiled in 2010 to mark its 25th anniversary (TERMCAT 2010a). According to the data, the Supervisory Council had held more than 500 meetings in which more than 6,500 terms had been normalised. In terms of developing terminological products, 515 works had been published, including compilations of terminology developed directly by TERMCAT terminologists with the assistance of experts, along with dictionaries developed by professionals, experts or researchers and later validated by the centre.

In addition to terminology work *per se*, TERMCAT has also contributed to establishing a working methodology for terminology and developed the technological resources to bring it to fruition. For this reason, the first period of the centre's operations were dedicated to drawing up all the necessary documentation, methodology and protocols; the terminological database was designed and implemented with an explicit plan for storing information from specialised vocabularies that were already published; the linguistic foundations for the normalisation of new terms were established; numerous collections of publications were launched; and through contacts with the outside world, a network of expert consultants was established, some of whom are still on the centre's technical committees today. The centre also opened its doors to the outside world by sharing its efforts and their specific value in Catalonia through participation in conferences and seminars.

During its more than 30 years of operation, TERMCAT has updated and expanded these resources. One such example is the centre's guide to developing systematic terminological studies in a specialised field, the first version of which was published in 1990 with the title of *Metodologia del treball terminològic*, followed by a revised version in 2010, called *El diccionari terminològic*. Other methodological, terminological and linguistic criteria have also been published.

Despite this, a vast amount of the terminological resources currently available for the Catalan language were not exclusively developed or promoted by TERMCAT.

Indeed, plenty of other organisations, institutions, companies and private individuals have participated in this activity out of a sense of duty to the language and its normalisation. As mentioned above, university and administrative language services are the most paradigmatic examples of this.

Technology at the service of terminology has also been an arena that has been particularly cultivated by Catalan researchers, either in the guise of digital terminology resources (digital dictionaries, terminology databases, translation memories) or resources directly applied to terms (terminology managers, term extractors), as well as tools that facilitate any of the tasks involved in the process of developing specialised dictionaries (creation, exploration and management systems of linguistic corpora, lemmatisers and labellers, parallelisers, analysers, etc.). Because of its importance, we should mention the TERMINUS work platform (<http://terminus.iula.upf.edu>) developed by the IULATERM research group at Universitat Pompeu Fabra (Barcelona), which allows all the phases in the development of a terminological project to be carried out automatically. Another project worth highlighting is an automatic tool to track normalised terminology, ESTEN (<http://esten.iula.upf.edu>), which allows quantitative and qualitative studies to be performed on the degree of acceptance of institutionally normalised terms dating back to 1985.

In addition, the restructuring of degrees in translation and interpreting at Spanish universities paved the way for teaching terminology as a compulsory subject as part of the core credits in the training of future professionals in this field. This was the first time that teaching of terminology became an official part of university-level studies. Terminology is officially taught in all the universities in the Catalan-speaking regions that have degree programmes in translation (Universitat Autònoma de Barcelona, Universitat Pompeu Fabra, Universitat de Vic, Universitat Jaume I de Castelló, Universitat de València and Universitat d'Alacant). It is also taught indirectly in courses on applied linguistics at philological faculties and in some documentation programmes, and in post-graduate programmes involving translation technologies, automatic language processing and other specialities (in this case associated with technical writing). Finally, terminology is also taught in all the language courses which specifically train teachers of Spanish and Catalan as a foreign language, and in foreign language classes for students in different specialities who want to acquire language competency in a professional domain. Therefore, given this context, we can claim that terminology has a strong presence at Catalan universities, and language professionals are thus increasingly aware of the existence of terminology and its role in representing and communicating knowledge, as well as in promoting the use of the language.

6 Terminology and neology

Catalan neological activity is closely associated with resolving terminology. The need to update the specialised lexicon in Catalan in all thematic fields means that TERM-CAT is constantly involved in neological activity. All the new terms can be found in its Neoloteca (<http://www.termcat.cat/neoloteca>).

Alongside TERMCAT's occasional neological activity, which consists of searching for an appropriate Catalan denomination for new terms, systematic neological activity in Catalan involves compiling the neologisms that appear in oral and written discourse. This task is performed by the Neology Observatory (OBNEO).

The OBNEO was established at the Universitat de Barcelona in 1989 and has been housed at Universitat Pompeu Fabra since 1994. Since 1989, it has been compiling neologisms that appear in the Catalan press but are not found in dictionaries. The bank of neologisms (BOBNEO) currently contains around 150,000 units, many of which have already been included in subsequent editions of the *Diccionari de la llengua catalana*.

The OBNEO has evolved considerably over the years:

The first stage (1989–1994) of its development focused on establishing the working methodology and training the people working on the project, most of them students and recent graduates of the Universitat de Barcelona.

The second stage (1994–1999) started automating the work of extracting neologisms from the written press and introduced new sources where neologisms could be found, namely the oral press and informal publications.

In the third stage (2000–2008), regionally distributed neology networks started being created, including Neological Antennas, the Spanish neology network which compiles units from all over Spain and Latin America; NEOROC, the neology network in peninsular Spain; NEOROM, a network which encompasses neology in the Romance languages; and more recently NEOXOC, a neology network for the different varieties of Catalan. During this stage, too, the automated structure which allows the entire chain of work to be carried out online was created. This period culminated with Barcelona hosting the First International Congress on Neology in Romance Languages (CINEO), which is held every three years in a country where a Romance language is spoken (Barcelona 2008, São Paulo 2011, Salamanca 2015, and Lyon 2018).

7 Summary and conclusions

As reflected in this chapter, the development of terminology as a recognised subject has been a long, slow process, and we cannot yet say that it has become definitively entrenched among the different fields of knowledge. By contrast, recognition of terminological practice linked to a variety of professional sectors, especially translation, has spread as globalisation has taken hold and the market of multilingualism

has grown stronger. All of the prestigious international organisations have their own dedicated translation services and tend to have complementary terminological support. In less important organisations, the translators themselves are in charge of resolving terminology. Yet somehow, terminology is always present.

There are many factors that might explain why it is so difficult to consolidate terminology as a new field of knowledge. First, terminology is based on practice, and applied disciplines generally have a harder time establishing themselves as fields of study in their own right, no matter how much theory there is behind them. Secondly, rather than being merely an interdisciplinary field, terminology is a transversal field, and this transversality implies a total shift in perspective from the traditional conception of how fields of knowledge are established over time. Finally, the normative origins of terminology, so closely tied to international normalisation, have also served to distance it from scholarly fields. For many years, prescriptive terminology was imposed as the dominant example, thus depriving it of the scientific (data-oriented) approach that characterised other fields of knowledge.

Catalan terminology has played and continues to play a very prominent role on the international scene. Judging from the recognition it has garnered, the success of the Catalan terminology model is an undeniable fact. This entails assessing not only its results but also, and even more importantly, its processes. Today it serves as a model for communities or countries in similar situations (minority or minoritised languages) or with similar goals (updating and promoting the language in specialised usage).

Catalan terminology's theoretical contribution to international knowledge has been extraordinary. The universal validity of Wüster's principles was first questioned by Catalonia, not Quebec. Catalonia also pointed out the contradiction between socio-terminological practices (although this term is not from our field) and the principles of Eugen Wüster's General Theory of Terminology.

However, Catalonia's contribution has gone even further than its material contributions; it has proved itself capable of developing a theoretical framework which, while not denying the validity of Wüster's conception in certain social and economic contexts, describes and explains terminology as an object of complex knowledge. This is based on three essential principles: the *polyhedricity principle* of all objects of knowledge, according to which every object of knowledge has several facets of description; the *multiple approach principle* to complex objects, which explains that the description of a polyhedral object can be approached from different facets; and the *non-contradiction principle*, which claims that the various approaches to an object must have the same conception of the object (Cabr  2003). The *Door Model* (Cabr  2000) allows the first two principles to be visualised.

8 Bibliography

- Alves, Ieda Maria/Simões Pereira, Eliane (edd.) (2015), *Neologia das línguas românicas*, São Paulo, Humanitas, CAPES.
- Cabré, M. Teresa (1999), *La terminología. Representación y comunicación*, Barcelona, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- Cabré, M. Teresa (2000), *Terminologie et linguistique: la théorie des portes*, Terminologie et diversité culturelle, Terminologies nouvelles 21, 10–15.
- Cabré, M. Teresa (2003), *Theories of Terminology. Their Description, Prescription and Explanation*, Terminology 9/2, 163–200.
- Cabré, M. Teresa (2004), *Terminologia i llengua catalana: funció social i cooperació*, in: Josep Massot i Muntaner (ed.), *Miscel·lània Joan Veny*, vol. 4, Barcelona, Publicacions de l'Abadia de Montserrat, 181–208.
- Cabré, M. Teresa (2009), *El context català, una àrea de concentració de coneixement i activitat terminològica*, Terminàlia 0, 8–14.
- Cabré, M. Teresa (2011), *La neología y los neologismos: reflexiones teóricas y cuestiones aplicadas*, in: María Eugenia Vázquez Laslop/Klaus Zimmermann/Francisco Segovia (edd.), *De la lengua por sólo la extrañeza: estudios de lexicología, norma lingüística, historia y literatura en homenaje a Luis Fernando Lara*, México, D.F., Colegio de México, 465–487.
- Cabré, M. Teresa/Domènech, Ona/Estopà, Rosa (edd.) (2014), *Mots nous en català/New Words in Catalan. Una panoràmica geolèctal/A Diatopic View*, Amsterdam/Philadelphia, Benjamins.
- Cabré, M. Teresa/Estopà, Rosa (edd.) (2009), *Les paraules noves: criteris per detectar i mesurar els neologismes*, Vic/Barcelona, Eumo/Universitat Pompeu Fabra.
- Cabré, M. Teresa/Estopà, Rosa/Vargas-Sierra, Chelo (edd.) (2012), *Neology in Specialized Communication*, Terminology 18/1, 1–8.
- Cabré, M. Teresa/Freixa, Judit/Solé, Elisabet (2000), *La neologia en el tombant de segle*, Barcelona, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- Cabré, M. Teresa/Freixa, Judit/Solé, Elisabet (edd.) (2008), *Lèxic i neologia*, Barcelona, Observatori de Neologia, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- Cabré, M. Teresa, et al. (edd.) (2010), *Actes del I Congrés Internacional de Neologia de les Llengües Romàniques*, Barcelona, Institut Universitari de Lingüística Aplicada, Universitat Pompeu Fabra.
- Cabré, M. Teresa, et al. (2015), *La neologia lèxica: darreres recerques sobre el català*, Caplletra 59, 123–124.
- DIEC = Institut d'Estudis Catalans (2007), *Diccionari de la llengua catalana*, Barcelona, Institut d'Estudis Catalans.
- Freixa, Judit/Bernal, Elisenda/Cabré, M. Teresa (edd.) (2015), *La neologia lèxica catalana*, Barcelona, Institut d'Estudis Catalans, Secció Filològica.
- Observatori de Neologia (1998), *Diccionari de paraules noves: Neologismes recollits a la premsa*, Barcelona, Universitat Pompeu Fabra/Enciclopèdia Catalana.
- Observatori de Neologia (2004), *Llengua catalana i neologia*, Barcelona, Meteora.
- Observatori de Neologia (ed.) (2016), *Mots d'avui, mots de demà*, Barcelona, Institut de Lingüística Aplicada, Universitat Pompeu Fabra.
- Sánchez Ibáñez, Miguel, et al. (edd.) (2017), *La renovación léxica en las lenguas románicas. Proyectos y perspectivas*, Murcia, Editum, Universidad de Murcia.
- Secretaria de Política Lingüística (2002), *Decret 107/1987, de 13 de març (DOGC 827, de 10 d'abril), pel qual es regula l'ús de les llengües oficials per part de l'Administració de la Generalitat de Catalunya. Els articles 5, 7 i 16 estan redactats d'acord amb el Decret 254/1987, de 4 d'agost (DOGC 885, de 2 de setembre). Afegit l'article 24 pel Decret 161/2002, d'11 de juny, i els apartats 2*

i 3 de l'article 2 pel Decret 162/2002, de 28 de maig (DOGC 3660, de 19 de juny), Diari Oficial de la Generalitat de Catalunya 827.

Societat Catalana de Terminologia (2011), *Estatuts de la Societat Catalana de Terminologia*, Barcelona, Societat Catalana de Terminologia.

TERMCAT, Centre de Terminologia (1990), *Metodologia del treball terminològic*, Barcelona, Generalitat de Catalunya, Departament de Cultura.

TERMCAT, Centre de Terminologia (2006), *La normalització terminològica en català: criteris i termes 1986–2004*, Barcelona, Publicacions de l'Abadia de Montserrat.

TERMCAT, Centre de Terminologia (2010a), *25 anys, 25 fites*, Barcelona, TERMCAT, Centre de Terminologia.

TERMCAT, Centre de Terminologia (2010b), *El diccionari terminològic*, Vic/Barcelona, Eumo Editorial/TERMCAT, Centre de Terminologia.

Emili Boix-Fuster and Kathryn A. Woolard

23 Language Ideologies in Society

Abstract: This chapter briefly defines the concept of language ideologies and identifies several principal ideological forms and processes. The appearance of such ideologies both in public debates and in everyday life in Catalan-speaking territories is reviewed, with attention to changes and differences across time and space. The role of language ideologies in establishing linguistic authority, and the basis of such authority in ideologies of linguistic authenticity or anonymity in Catalan-speaking areas, are discussed. The review considers ideological mechanisms of linguistic and social differentiation, focusing in particular on the iconization of Catalan, erasure of forms and patterns of use that do not fit these iconizations, and the recursive particularism that ideologically fragments the unity of the language. The final sections of the review address language professionals' and elites' stances toward bilingualism in contrast to plurilingualism, and new ideological bases for the defense of Catalan that have been emerging in recent years.

Keywords: authenticity, anonymity, iconization, catalanization, recursive particularism, bilingualism, Catalan-speaking territories

1 Introduction

Language ideologies are morally and politically loaded representations of the structure and use of languages in the social world. They occur in ingrained, implicit sensibilities as well as in explicit discourses. Such ideologies can reflexively shape social as well as linguistic processes while purportedly simply representing them, and thus they are an indispensable facet of sociolinguistic research (Silverstein 1979; Woolard 1998). In Catalan-speaking territories as elsewhere, conflicting and consequential ideological representations are produced by elites, community members, and language activists and professionals (including the present authors). In this chapter, we review just some of the most salient facets of language ideologies in Catalan-speaking territories.

2 Linguistic authority

The modern Western world is dominated by two ideologies of linguistic authority that have been called “authenticity” and “anonymity” (Woolard 2016). As a legacy of Romanticism, “authenticity” confers value on a linguistic form taken to express the

<https://doi.org/10.1515/9783110450408-031>

true nature of a group or an individual, and to be rooted in that particular social experience. In contrast, “anonymity” is the value of languages alleged to be universally available to all and not reflective of a particular social position.

These two ideologies have both legitimized and constrained the Catalan language at least since the 19th-century *Renaixença* (≈13 *Renaixença*), when early Catalanism first cast the language as the *sine qua non* of the Catalan nation (see ≈1 Languages, Cultures, Nations: A History of Europe). This brought a nostalgic reappraisal of past times and especially of a medieval Catalan “Golden Age”, and of the language as an index of the historical depth of Catalan authenticity. Particularly in Catalonia, sectors of the Catholic Church contributed to the use of Catalan on the premise that it was the natural language of the people (Torras i Bages 1981, 1892). In the 20th-century transition to the post-Franco Spanish state as well, the defense of Catalan as the *llengua pròpia* (‘own language’) of Catalan territories drew its legitimacy from this ideology of authenticity (Sinner/Wieland 2008; Solé i Durany 1996; Süselbeck 2008; Wurl 2011).

By the turn of the 20th century the main dynamic of Catalan culture affirmed modernity with a nationalist emphasis, still focusing on language as an icon of the nation (Prat de la Riba 1906). Catalanism enjoyed the support of sectors of the bourgeoisie. This allowed both language reform (Costa 2009) and some public, official use of Catalan (Galí 1979). The language reformer Pompeu Fabra saw Catalan as a symbol of the nation that needed decastilianization and purification. His ideology of the language paralleled nationalist values: a high degree of internal cohesion coupled with sharp distinction from external groups and languages (Lamuela/Murgades 1984). Fabra’s reform succeeded more in Catalonia, but less so in Valencia and the Balearics.

Most Catalanist intellectuals have construed the language as the core of the Catalan nation, and defenses of Catalan have been situated within this same ideology of authenticity (Ortega 2016). The loss of this icon of identity is feared as an existential threat, as in this argument from Joan Fuster, a leading Valencian essayist: ‘Deposit of centuries, heritage of *convivència*, framework of culture, the language is both axis and continuity for the life of a people, as a people. If the language is torn, bastardized or lost, the society is broken’ (Fuster 1963, lx–lxi). The Catalan linguist Joan Solà similarly stressed the significance of language for the inner self: ‘A language, for those of us for whom it is a patrimony, is as inseparable from ourselves as is our blood, as is any aspect of our personality, the name we bear, or the color of our skin’ (Solà 2011, 62). Accordingly, many Catalan activists experience ‘language pain’ over the loss they see of both its use and structure (Larreula 2002; Vidal 2015).

The rest of the population, in contrast, is generally more indifferent, and young people in particular frame language as a matter of personal choice rather than a political principle of community (Flors Mas/Vila i Moreno 2014; Pujolar 2001; Woolard 2016). Nor are all Catalan sociolinguistic observers so pessimistic (see, e.g., Baulenas 2004; Pueyo 2007; Querol Puig/Strubell i Trueta 2009). But speakers do not often

question the habits that privilege Castilian on the basis of its “misrecognized” ideological anonymity, in Bourdieu’s term. Accommodation to Castilian in interethnic encounters seems ingrained in the bodily dispositions of the Catalan linguistic habitus (Bastardas i Boada 1995; Bourdieu 1982).

3 Authenticity and anonymity as constraints on catalanization

The Francoist dictatorship prohibited the public use of Catalan and blocked the smooth linguistic integration of the massive immigrant population that arrived in Catalan-speaking lands, especially during the 1960s (Domingo 2014). The maintenance of the language in Catalonia owed to its distinctive socio-economic structure and to concomitant social connotations. The Catalan language was associated with the middle class, who used it at home and in public even during harsh dictatorial periods (Boix-Fuster/Paradís 2015), with the “Nova Cançó” movement as an example (Cubeles/Ruiz 2010). However, longstanding Castilian hegemony and the linguistic repression of the twentieth century left deep ideological scars. Diglossic ideologies (Marfany 2008) were widespread among the population, making only Castilian (and French in Northern Catalonia even more) seem suitable for formal domains, and relegating Catalan to private – if any – spheres.

In post-Franco Spain, Catalan is recognized officially, albeit on an unequal footing with Castilian, and is designated as Catalonia’s *llengua pròpia* (Solé i Durany 1996). This authenticity-based concept supports its public use but draws criticism as illiberal not only from Castilianist opponents (Santamaría 1999), but also from iconoclastic defenders of Catalan who base their defense on contemporary democratic choice (Branchadell 1997).

Under these conditions, the authenticity value of Catalan has been both a motive for and a constraint on its acquisition as a second language (Pujolar/González 2013; Vila i Moreno 2004a). Speaking Catalan has been taken as essential to being authentically Catalan, so in turn, it is often assumed that only Catalans speak Catalan, and that it is off-limits to outsiders and foreigners (Sabaté i Dalmau 2016). This is reinforced by Catalan speakers’ habit of switching to Castilian with those they take not to be habitual Catalan speakers. Early campaigns to “normalize” Catalan promoted passively bilingual conversations, with interlocutors each maintaining their “own” language, in keeping with the authenticity frame (Boix-Fuster/Melià/Montoya 2011; Suay/Sangenís 2010). Recent governmental campaigns continue to encourage Catalan speakers to combat the habit of automatically switching to Castilian to address perceived foreigners.

In policy conflicts in recent decades, Castilianist proponents often attribute to Castilian the value of anonymity, legitimating it as the “common language” seemingly

equally available for use by all, regardless of origins. They represent Catalan, in contrast, as a restrictive sign of a particular identity unavailable to non-native speakers (Lodares 2000; Vargas Llosa et al. 2008; for critical commentary on this view and more examples, see Argenter 2007; Moreno Cabrera 2015; Ortega 2016). Castilian is further cast as a vehicle of cosmopolitanism and a superior economic instrument, while Catalan is depicted as parochial and irrational in the marketplace (Boix-Fuster 1994; DiGiacomo 1999). These views are channelled by various Spanish intellectuals and by powerful Spanish political parties, such as the Partido Popular and Ciudadanos (Pericay 2007), which buttress Castilian within discourses of free choice and free markets (Duane 2017). In a notable exception to this Castilianist bent, the Spanish linguist Juan Carlos Moreno Cabrera (2015) has critically analyzed numerous examples of such hegemonic Spanish linguistic ideologies.

The ideology of Catalan authenticity is still robust, as seen in the recent Koiné Group manifesto (Grup Koiné 2016), which sounded an alarm about bilingualism as a perceived imminent threat to Catalan (Murgades 2016). Nonetheless, anonymity may be essential to the goal of “normalizing” Catalan as an unmarked vehicle of communication by all kinds of speakers in all domains (Castillo 2006; Flaquer 1996; McRoberts 2001; Vallverdú 1990). The main challenge to the societal spread of Catalan is to make it a public language used beyond the boundaries of the native Catalan community, reaching the rich variety of immigrants, from Latin Americans to Rumanian or Pakistani new settlers. Some recent Catalan advocacy has shifted away from the constrained ideological foundation of authenticity and toward anonymity, a stance sometimes lauded, sometimes criticized as “post-national” (Carod-Rovira 2012; Pujolar 2007a; Soler 2013). Organizations such as *Plataforma per la llengua* have devised innovative campaigns such as “Catalan, common language”, inviting non-native speakers to join the Catalan linguistic community (Biosca 2009; Plataforma per la Llengua 2010; Ponsa 2016; Riera Gil 2013; Vila i Ros 2007). Similarly, *Veu Pròpia* (‘Our Own Voice’), a civic organization of L2 Catalan speakers, reformulated as legitimate a personal relation to the language based on choice rather than native speaker roots. This is bolstered by research evidence of a gradual shift among younger generations that de-ethnicizes the social meaning of the language (Newman/Trenchs-Parera/Ng 2008; Pujolar/González 2013; Woolard 2016). But despite such efforts and trends, there are still covert biases among defenders of Catalan toward conceiving of the language as owned by native speakers (Pujolar 2007b).

4 Linguistic differentiation

Distinct languages are produced by continual processes of differentiation and consolidation by professionals and speakers (Bakhtin 1981; Irvine/Gal 2000). These processes create typifications of languages and of the speakers with which they are associated, drawing on three ideological techniques identified by Irvine and Gal:

iconization interprets a linguistic form as an image of the essential nature of its speakers; *erasure* overlooks or eliminates forms and speakers that do not fit that image; *fractal recursivity* projects these typifications onto multiple levels of social life.

The most apparent iconization in the Catalan case is the link of the language to Catalan identity. Nonetheless, movements against the repressive Franco régime tempered this. Sectors of political opposition as well as of the Church proclaimed ‘everybody who lives and works in Catalonia is Catalan’ (Candel 1964; Pujol 1976), erasing the boundaries between natives and newcomers. This optimistic ideology was reflected in the first language campaign in 1980: ‘Catalan is everybody’s business’ (Boix-Fuster/Melià/Montoya 2011).

4.1 Iconization: Language, class, and gender

In the past, an influential upper class in Catalonia became castilianized and stigmatized Catalan as uncouth. This stratum persists, and Catalan has often been characterized on the basis of patterns of use and non-use as lacking prestige. However, Catalan has come to be associated with the middle and upper middle class in Catalonia, leading to its being viewed there (less so elsewhere) as a vehicle of social promotion for working and lower middle class speakers of non-Catalan origins. Studies of covert linguistic attitudes and overt discourses there show that at the level of individual consciousness, Catalan endows a speaker with considerable prestige (Woolard 1989), although this effect may be complicated by recent immigration and social change (Newman/Trenchs-Parera/Ng 2008). In Valencia, in contrast, Catalan is stigmatized as rustic (González Martínez 2010). Moreover, working class L1 Catalan speakers may actually be more disadvantaged by bureaucratic requirements of linguistic mastery for some employment than is generally recognized (Frekko 2013). Gender is intertwined with social class in modern societies, so it is not surprising that an incipient gendering of Catalan as feminine or effeminate, and therefore ill-suited for the expression of a heteronormative working class masculinity, has been identified in urban areas (Frekko 2009; Pujolar 2001).

Studies of language attitudes show that in Andorra as well as Catalonia, the subjective ethnolinguistic vitality of Catalan is generally high. That is, community members judge its prestige and usefulness as strong compared to others (Viladot 1993). Atkinson, however, argues that the concept of ethnolinguistic vitality is difficult to apply to the Catalan case (Atkinson 2000). Because of its value for social advancement, most teachers, even those of non-Catalan origin, evaluate positively the catalanization of the education system (Baron 2015), where Catalan is formally dominant (Vila i Moreno 2004b). In interviews of linguistically mixed families in contemporary Catalonia, attitudes toward Catalan are mainly positive, and Castilian-speaking spouses even come to speak Catalan to their children (Boix-Fuster/Paradís 2015).

Simultaneously, though, most respondents feel strongly that the acquisition of Spanish is necessary, and more recently that English is also indispensable.

The perceived relation between the Catalan language and elevated social status goes beyond this predictable association (known technically as indexicality), to an iconization of the language itself in class-cultural terms. Catalan has come to be heard by many in Catalonia, particularly young people, as a “refined” way of speaking, in contrast to Castilian, which is stereotypically perceived as “coarse” (Flors Mas/Vila i Moreno 2014; Woolard 2016). Even specific morphosyntactic forms come to be ideological icons. Most famously, the distinctive system of weak pronouns with its prescriptive complexity is an icon of the allegedly true, (hyper)refined nature of Catalan and Catalan identity (Frekko 2009; Woolard/Ribot Bencomo/Soler Carbonell 2014). Iconization is always accompanied by erasure, and in the case of Catalan in contemporary Catalonia, vulgar, rough, and ludic registers of the language are not only overlooked but may even have been stripped from the functions of Catalan, particularly for young speakers, partly as an unintended consequence of educational and mass media language policies (Frekko 2009). In adolescence, young people may take up popular culture in which Castilian predominates, and some may favor it to express transgressive stances (Bretxa/Vila i Moreno 2012). However, the novels of Ramon Solsona, for example, show that Catalan can still express coarse messages, and non-normative colloquialisms and mixed languages also appear in recent literature (Boix-Fuster/Riba 2017).

Linguistic norms have been a recurrent ideological battlefield (Ferrando/Nicolás 2011; Jané-Lligé 2017; Vallverdú 1968). Purist linguistic stances are pervasive and have produced bitter public battles, such as one between so-called “Light Catalan” and “Heavy Catalan” norms for media use in the late 1980s and early 1990s (Pericay/Toutain 1986; Prats/Rafanell/Rossich 1991). In one of few sociolinguistic studies of Catalan linguistic purism and conservatism, Frekko (2009) argues that language and media professionals reduce their vision of the language to a single standard register and ideologically erase other varieties, making it appear more fragile than it actually is by adhering to ideologies of authenticity and the monolingual nation, and to rigid ideological boundaries between languages. This fragility was apparent when, in late 2016, the Academy of the Catalan Language (Philology Section of the Institut d’Estudis Catalans) made public some changes in orthographic and grammatical norms, triggering a conservative uproar, especially among literati and linguists (Murgades 2016).

4.2 Recursive particularism

Valencia is the second major center of Catalan culture, but a distinctive political-economic history has eroded the language in its most important urban centers, where it is now marginal. Hence Valencian images of Catalan tend to be more negative (Cuenca 2003). Some speakers despise their own language, in what has been called

linguistic ‘self-hatred’ (Ninyoles 1972). (For more nuanced later findings, see Cases-noves 2010.)

The iconic relation of language to identity is recursively reproduced in a particularistic struggle over the name of the linguistic form spoken in Valencia (Pradilla Cardona 1999). Conservative-led sectors of Valencian society refuse to see their linguistic variety as part of a common Catalan language, despite overwhelming endorsement of such unity by linguists. The motto of this particularism is: “No mos fareu catalans” (‘You will not make us Catalans’). This reveals a reading of the boundaries of a named language as iconic of the boundaries of a distinct people. From this perspective, labeling the Valencian linguistic variety as Catalan supports alleged imperialistic pretensions from Catalonia, while the label Valencian is an iconic barrier to cultural subordination. Thus, the popular and official name of the language there is Valencian, and since 2001 there has been an independent normative linguistic authority for it, the Acadèmia Valenciana de la Llengua). Similar particularism challenging the unity of the Catalan language appears in Majorca, in a weaker form known as *gonellisme* (Moll 1972). In both communities, these defenses of localist authenticities are linked to conservative political parties that naturalize a monolithic Spanish hegemony while subjecting Catalan to a fractally recursive ideology of fragmentation (Bibiloni 1999; Duane 2017).

In Northern Catalonia as in Valencia, Catalan is a denigrated variety. The homogenizing ideology of the French Republic, stemming from the Jacobin French Revolution (Baylac-Ferrer 2014; Schlieben-Lange 1996), has held that all non-French languages of France are backward, parochial, and inimical to both human rights and human progress. This ideology has been so hegemonic that Catalan has been almost totally abandoned as a first language in France.

5 Bilingualism vs plurilingualism

There has been a double resistance to bilingualism in Catalan intellectual milieux, which can puzzle defenders of bilingualism among minoritized speakers elsewhere. From a psychological perspective, bilingualism was seen as a threat to an authentic self, as in this negative evaluation by Antoni Badia i Margarit, a leading Romance linguist: ‘Bilingualism in individuals is an unstable situation, that leads inevitably to monolingualism [...]. And it is good that it is unstable. If language is the most intimate expression of personality, the possibility of expressing oneself equally in two languages comes to affect the unity of the personality. And it goes without saying that this is very serious from a psychological perspective’ (Badia i Margarit 1966, 97). From a sociological point of view, the pioneering Valencian sociolinguist Lluís V. Aracil denounced what he called the *myth of bilingualism*, which holds that Valencia has always been *naturally* bilingual. He analyzed bilingualism as a mask for inexorable language shift to Castilian (Aracil 1966), a shift that is still occurring (Montoya/Mas 2011). Aracil’s critique of bilingualism became a generalized orthodox position in

Catalan sociolinguistics, although it has been questioned more recently by professional observers sympathetic to Catalan (Atkinson 1998; Branchadell 2012; Ortega 2015).

Most Catalan language activists are still wary of “bilingualism”, but “multilingualism” and “polylingualism” are now widely accepted and used to challenge the ideologies of both anonymity and authenticity. Linguistic cosmopolitanism is promoted by some sectors and associated not with one language or the other, but rather with multiplicity and fluidity as tools for social harmony, in contrast to “linguistic parochialism”, which clings to monolingualism as an index of personal identity (Trenchs-Parera/Larrea Mendizabal/Newman 2014).

Recently, an “ecological” ideology has emerged in Catalanist attempts to nuance the ideology of authenticity in making a case for Catalan as the essential language of the Catalan-speaking territories. In a universalist endeavour built on a clear Catalan base, the *Universal Declaration of Linguistic Rights* epitomizes this perspective and its foundation in authenticity ideology; e.g. Article 7: “All languages are the expression of a collective identity and of a distinct way of perceiving and describing reality and must therefore be able to enjoy the conditions required for their development in all functions” (Universal Declaration of Linguistic Rights. Follow-up Committee 1996). Albert Bastardas i Boada (2014) argues for the preservation of an ecology of linguistic and social diversity, to avoid the poverty, anomie, and loss of cultural wisdom that result from disorganisation of the traditional subsistence ecosystem. He follows the biologist Ramon Folch in advocating “linguistic sustainability”, a “gradual transformation from the current model of the linguistic organisation of the human species, [...] whose objective would be to avert the implication that collective bilingualism or polyglotism [...] requires the abandonment by different cultural groups of their own languages” (Bastardas i Boada 2014, 138, based on Reales 1999).

Similarly, the leaders of GELA (Group in Defense of Threatened Languages) at the University of Barcelona criticize state policies as linguicidal and defend the protection of diverse languages through mutual recognition (Junyent 1998). They embrace on principle all languages in Catalonia and advocate Catalan as the language of common interchange, as the most historically rooted and least imbricated with hierarchy and colonialism (Comellas 2009). Their stated hope is that such active mutual respect for subordinated languages might trigger a “mirror effect”, in which migrants positively re-evaluate their own formerly despised native language.

6 Conclusion

There is no clearly dominant language ideology in the Catalan-speaking territories. Within Catalonia alone, for instance, the most influential newspaper, *La Vanguardia*, defends the maintenance of Spanish-medium cultural production in a bilingual Catalonia, against autonomous government policies that support the economically more fragile cultural production in Catalan. Across the Catalan-speaking territories,

as we have suggested above, there are even more marked variations in elite as well as everyday ideological positionings in relation to Catalan. Moreover, significant new formulations have been emerging that challenge the traditional ideological foundations of linguistic practices in the Catalan-speaking territories. The ideological struggle over language is alive and overt in the public arena, particularly at the intersection of the Spanish state with the autonomous communities. It is at the same time at work, even though relatively covertly, in the everyday life of speakers.

7 Bibliography

- Aracil, Lluís V. (1966), *Bilingualism as a Myth*, *Interamerican Review* 2/4, 521–533.
- Argenter, Joan A. (2007), *Els discursos sobre la situació lingüística a Catalunya. Diagnosi i prospectiva: de la monotonía a la polifonia*, in: Barbara Czernilofsky et al. (edd.), *El discurs sociolingüístic actual català i occità. Lo discurs sociolingüístic actual catalan e occitan*, Viena, Praesens, 135–147.
- Atkinson, David (1998), *Normalization: Integration or Assimilation? A Response to Miquel Strubell*, *Current Issues in Language & Society* 5/3, 210–214.
- Atkinson, David (2000), *Minoritisation, Identity and Ethnolinguistic Vitality in Catalonia*, *Journal of Multilingual and Multicultural Development* 21/3, 185–197.
- Badia i Margarit, Antoni M. (1966), *La integració idiomàtica i cultural dels immigrants. Reflexions, fets i plans*, *Qüestions de vida cristiana* 31, 91–103.
- Bakhtin, Mikhail M. (1981), *The Dialogic Imagination*, ed. Michael Holquist, Austin, University of Texas Press.
- Baron, Joan (2015), *Ideologies lingüístiques dels mestres de les promocions del pla 67 (1967–72) i del Pla Pilot de Bolonya (2004–09) envers les llengües de l'educació primària de Catalunya*, Ph.D. thesis, Barcelona, University of Barcelona.
- Bastardas i Boada, Albert (1995), *Language Management and Language Behavior Change: Policies and Social Persistence*, *Catalan Review: International Journal of Catalan Culture* 9/2, 15–38.
- Bastardas i Boada, Albert (2014), *Linguistic Sustainability for a Multilingual Humanity*, *Sustainable Multilingualism* 5, 134–163.
- Baulenas, Lluís-Anton (2004), *El català no morirà. Un moment decisiu per al futur de la llengua*, Barcelona, Edicions 62.
- Baylac-Ferrer, Alà (2014), *La ideologia lingüística francesa i les llengües regionals avui. El cas del català a la Catalunya del Nord*, *Treballs de Sociolingüística Catalana* 24, 57–87.
- Bibiloni, Gabriel (1999), *La llengua estàndard en una situació de normalitat*, in: Miquel Strubell i Trueta et al. (edd.), *El català: Un debat a finals del segle XX. De la normalització i la normalitat*, Barcelona, La Busca, 63–70.
- Biosca, Marc (2009), *“El català, llengua comuna” torna a sortir al carrer*, *La Corbella* 15, 4.
- Boix-Fuster, Emili (1994), *Ideologies lingüístiques de les generacions joves a Barcelona*, in: Gabinet Municipal de Normalització Lingüística (ed.), *La política lingüística a l'Estat espanyol: balanç i perspectives*, Alcoi, Ajuntament d'Alcoi, 75–101.
- Boix-Fuster, Emili/Melià, Joan/Montoya, Brauli (2011), *Policies Promoting the Use of Catalan in Oral Communications and to Improve Attitudes Towards the Language*, in: Miquel Strubell/Emili Boix-Fuster (edd.), *Democratic Policies for Language Revitalization: The Case of Catalan*, New York, Palgrave Macmillan, 150–181.
- Boix-Fuster, Emili/Paradís, Anna (2015), *Ideologies and Trajectories of “New Speakers” in Bilingual Families in Catalonia*, *Revista de Llengua i Dret* 63, 165–185.

- Boix-Fuster, Emili/Riba, Cèlia (2017), *Ramon Solsona: el plaer de la hibridació i el joc lingüístics*, Els Marges 111, 10–28.
- Bourdieu, Pierre (1982), *Ce que parler veut dire. L'économie des échanges linguistiques*, Paris, Fayard.
- Branchadell, Albert (1997), *Liberalisme i normalització lingüística*, Barcelona, Empúries.
- Branchadell, Albert (2012), *One Nation, One (Common) Language? Language and Nationalism in 21st Century Catalonia*, Recode Online Working Paper 10, <http://www.recode.info/wp-content/uploads/2013/02/Branchadell-Albert-2012-RECODE.pdf> (last accessed: 7.11.2018).
- Bretxa, Vanessa/Vila i Moreno, F. Xavier (2012), *Els canvis sociolingüístics en el pas de primària a secundària: el projecte RESOL a la ciutat de Mataró*, Treballs de Sociolingüística Catalana 22, 93–118.
- Candel, Francesc (1964), *Els altres catalans*, Barcelona, Edicions 62.
- Carod-Rovira, Josep-Lluís (2012), *Identitats i nació en construcció: El cas català*, Divèrsia 2, 98–122.
- Casesnoves, Raquel (2010), *Changing Linguistic Attitudes in Valencia: The Effects of Language Planning Measures*, Journal of Sociolinguistics 14/4, 477–500.
- Castillo, Mónica (2006), *Los discursos de los manuales de sociolingüística catalanes y la normalización: análisis desde los años 80 a la actualidad*, in: Mónica Castillo Lluch/Johannes Kabatek (edd.), *Las lenguas de España. Política lingüística, sociología del lenguaje e ideología desde la Transición hasta la actualidad*, Madrid/Frankfurt am Main, Iberoamericana/Vervuert, 223–240.
- Comellas, Pere (2009), *Les llengües immigrades: adversàries o aliades de les llengües minoritzades*, in: Carme Junyent (ed.), *Llengua i acollida*, Barcelona, Horsori, 23–41.
- Cubeles, Xavier/Ruiz, Isabel (2010), *La nova cançó. La veu d'un poble*, Barcelona, Generalitat de Catalunya.
- Cuenca, Maria Josep (2003), *El valencià és una llengua diferent?*, València, Tàndem.
- DiGiacomo, Susan (1999), *Language Ideological Debates in an Olympic City: Barcelona 1992–1996*, in: Jan Blommaert (ed.), *Language Ideological Debates*, Berlin/New York, Mouton de Gruyter, 105–142.
- Domingo, Andreu (2014), *Catalunya al mirall de la immigració. Demografia i identitat nacional*, Barcelona, L'Avenç.
- Duane, Lucas (2017), *Castilian Takes Backstage in the Balearic Islands: The Activation of Catalan Standardization Recursions in Facebook*, Journal of Linguistic Anthropology 27/1, 71–91. DOI:10.1111/jola.12141.
- Fabra, Pompeu (2009), *The Architect of Modern Catalan: Pompeu Fabra (1868–1948). Selected Writings*, ed. Joan Costa, Amsterdam/Philadelphia, Benjamins.
- Ferrando, Antoni/Nicolás, Miquel (2011), *Història de la llengua catalana*, Barcelona, UOC.
- Flaquer, Lluís (1996), *El català, ¿llengua pública o privada?*, Barcelona, Empúries.
- Flors Mas, Avel·li/Vila i Moreno, F. Xavier (2014), *Justificant les preferències. Com argumenten les opcions lingüístiques els adolescents catalans*, Treballs de Sociolingüística Catalana 24, 173–199.
- Frekko, Susan E. (2009), *“Normal” in Catalonia*, Language in Society 38, 1–23.
- Frekko, Susan E. (2013), *Legitimacy and Social Class in Catalan Language Education for Adults*, International Journal of Bilingual Education and Bilingualism 16/2, 164–176.
- Fuster, Joan (1963), *Introducció a la poesia de Salvador Espriu*, in: Salvador Espriu (ed.), *Obra poètica*, Barcelona, Albertí, xiii–lxxix.
- Galí, Alexandre (1979), *Història de les institucions i del moviment cultural a Catalunya 1900–1936. Llibre I: La llengua. Entitats defensores i propagadores*, Barcelona, Fundació Alexandre Galí.
- González Martínez, Juan (2010), *Actitudes lingüísticas en una comunidad rural: Els Ports (Castellón): datos de un cuestionario sociolingüístico*, Cultura, Lenguaje y Representación 8, 75–95.
- Grup Koiné (2016), *Per un veritable procés de normalització lingüística a la República Catalana*, <http://llenguarepublica.cat/prodsite/wp-content/uploads/2016/04/ManifestParanimf.pdf> (last accessed: 7.11.2018).

- Irvine, Judith T./Gal, Susan (2000), *Language Ideology and Linguistic Differentiation*, in: Paul Kroskrity (ed.), *Regimes of Language*, Santa Fe, New Mexico, School for American Research, 35–84.
- Jané-Lligé, Jordi (2017), *Models de llengua en les traduccions del segon franquisme*, in: Montserrat Barcardí/Pilar Godayol (edd.), *Traducció i franquisme*, Lleida, Punctum, 105–119.
- Junyent, Carme (1998), *Contra la planificació*, Barcelona, Empúries.
- Lamuela, Xavier/Murgades, Josep (1984), *Teoria de la llengua literària segons Fabra*, Barcelona, Quaderns Crema.
- Larreula, Enric (2002), *Dolor de llengua*, València, Edicions 3 i 4.
- Lodares, Juan Ramón (2000), *El paraíso políglota. Historias de lenguas en la España moderna contadas sin prejuicios*, Madrid, Taurus.
- Marfany, Joan-Lluís (2008), *Llengua, nació i diglòssia*, Barcelona, L'Avenç.
- McRoberts, Kenneth (2001), *The Catalan Language: The Politics of "Normalization"*, in: Kenneth McRoberts (ed.), *Catalonia. Nation Building Without A State*, New York, Oxford University Press, 139–160.
- Moll, Francesc de Borja (ed.) (1972), *Polèmica d'en Pep Gonella*, Palma de Mallorca, Moll.
- Montoya, Brauli/Mas, Antoni (2011), *La transmissió familiar del valencià*, València, Acadèmia Valenciana de la Llengua.
- Moreno Cabrera, Juan Carlos (2015), *Errores y horrores del españolismo lingüístico. Cinco vocales para conquistar el mundo*, Tafalla, Txalaparta.
- Murgades, Josep (2016), *Escrips sobre llengua*, Lleida, Pagès.
- Newman, Michael/Trenchs-Parera, Mireia/Ng, Shukhan (2008), *Normalizing Bilingualism: The Effects of the Catalan Linguistic Normalization Policy One Generation After*, *Journal of Sociolinguistics* 12/3, 306–333.
- Ninyoles, Rafael Lluís (1972), *Idioma y poder social*, Madrid, Tecnos.
- Ortega, Rudolf (2015), *El bilingüisme muta*, *El País*, http://cat.elpais.com/cat/2015/04/05/cultura/1428264483_338515.html (last accessed: 7.11.2018).
- Ortega, Rudolf (2016), *En defensa de la llengua. Discursos, manifestos i textos decisius en la història del català*, Barcelona, Angle Editorial.
- Pericay, Xavier (2007), *¿Libertad o coacción? Políticas lingüísticas y nacionalismos en España*, Madrid, Fundación FAES.
- Pericay, Xavier/Toutain, Ferran (1986), *Verinosa llengua*, Barcelona, Empúries.
- Plataforma per la Llengua (2010), *El catalán también es mío/El català també és meu*, Barcelona, Fundació Vincle/Plataforma per la Llengua.
- Ponsa, Gemma (2016), *Amb el català, fem pinya*, Barcelona, Plataforma per la Llengua.
- Pradilla Cardona, Miquel Àngel (1999), *El secessionisme lingüístic valencià*, in: Miquel Àngel Pradilla Cardona (ed.), *La llengua catalana al tombant del mil·leni. Aproximació sociolingüística*, Barcelona, Empúries, 153–202.
- Prat de la Riba, Enric (1906), *La nacionalitat catalana*, Barcelona, L'Anuari de la Exportació.
- Prats, Modest/Rafanell, August/Rossich, Albert (edd.) (1991), *El futur de la llengua catalana*, Barcelona, Empúries.
- Pueyo, Miquel (2007), *El fantasma de la mort del català*, Barcelona, Proa.
- Pujol, Jordi (1976), *La immigració, problema i esperança de Catalunya*, Barcelona, Nova Terra.
- Pujolar, Joan (2001), *Gender, Heteroglossia and Power. A Sociolinguistic Study of Youth Culture*, Berlin/New York, Mouton de Gruyter.
- Pujolar, Joan (2007a), *Bilingualism and the Nation-State in the Post-National Era*, in: Monica Heller (ed.), *Bilingualism: A Social Approach*, New York, Palgrave, 71–95.
- Pujolar, Joan (2007b), *The Future of Catalan: Language Endangerment and Nationalist Discourses in Catalonia*, in: Alexandre Duchêne/Monica Heller (edd.), *Discourses of Endangerment*, London, Continuum, 121–148.

- Pujolar, Joan/Gonzàlez, Isaac (2013), *Linguistic “Mudes” and the De-Ethnicization of Language Choice in Catalonia*, International Journal of Bilingual Education and Bilingualism 16/2, 138–152.
- Querol Puig, Ernest/Strubell i Trueta, Miquel (2009), *Llengua i reivindicacions nacionals a Catalunya*, Barcelona, UOC.
- Reales, Lluís (1999), *Una conversa amb Ramon Folch*, Idees 2, 99–108.
- Riera Gil, Elvira (2013), *Sobre el concepte polític de llengua comuna: una aproximació teòrica i comparada*, Revista de Llengua i Dret 60, 91–110.
- Sabaté i Dalmau, Maria (2016), *El català dels locutors: pràctiques i ideologies lingüístiques de la migració en espais urbans castellanoparlants*, Treballs de Sociolingüística Catalana 26, 105–121.
- Santamaría, Antonio (1999), *Lengua propia, conducta impropia*, in: Antonio Santamaría (ed.), *Foro Babel: El nacionalismo y las lenguas de Cataluña*, Barcelona, Àltera, 181–197.
- Schlieben-Lange, Brigitte (1996), *Idéologie, révolution et uniformité de la langue*, Sprimont, Mardaga.
- Silverstein, Michael (1979), *Language Structure and Linguistic Ideology*, in: Paul R. Clyne/William Hanks/Carol L. Hofbauer (edd.), *The Elements: A Parasection on Linguistic Units and Levels*, Chicago, Chicago Linguistic Society, 193–247.
- Sinner, Carsten/Wieland, Katharina (2008), *El catalán hablado y problemas de la normalización de la lengua catalana: avances y obstáculos en la normalización*, in: Kirsten Süselbeck/Ulrike Mühl-schlegel/Peter Masson (edd.), *Lengua, nación e identidad. La regulación del plurilingüismo en España y América Latina*, Madrid/Frankfurt am Main, Iberoamericana/Vervuert, 131–164.
- Solà, Joan (2011), *La paraula*, in: Joan Solà (ed.), *L’última lliçó. Parlaments polítics i acadèmics*, Barcelona, Empúries, 57–65.
- Solé i Durany, Joan Ramon (1996), *El concepte de la llengua pròpia en el dret i la normalització de l’idioma a Catalunya*, Revista de Llengua i Dret 26, 95–120.
- Soler, Josep (2013), *The Anonymity of Catalan and the Authenticity of Estonian: Two Paths for the Development of Medium-Sized Languages*, International Journal of Bilingual Education and Bilingualism 16/2, 153–163.
- Suay, Ferran/Sangenís, Gemma (2010), *Sortir de l’armari lingüístic. Una guia de conducta en català*, Barcelona, Angle Editorial.
- Süselbeck, Kirsten (2008), *“Lengua”, “nación” e “identidad” en el discurso de la política lingüística de Cataluña*, in: Kirsten Süselbeck/Ulrike Mühl-schlegel/Peter Masson (edd.), *Lengua, nación e identidad: La regulación del plurilingüismo en España y América Latina*, Madrid/Frankfurt am Main, Iberoamericana/Vervuert, 165–186.
- Torras i Bages, Josep (1981, 1892), *La tradició catalana*, Barcelona, Edicions 62.
- Trenchs-Parera, Mireia/Larrea Mendizabal, Imanol/Newman, Michael (2014), *La normalització del cosmopolitisme lingüístic entre els joves del segle XXI? Una exploració de les ideologies lingüístiques a Catalunya*, Treballs de Sociolingüística Catalana 24, 281–302.
- Universal Declaration of Linguistic Rights. Follow-up Committee (1996), *Declaració universal de drets lingüístics. Déclaration universelle de droits linguistiques. Universal declaration of linguistic rights. Declaración universal de derechos lingüísticos*, Barcelona, Ciemen.
- Vallverdú, Francesc (1968), *L’escriptor català i el problema de la llengua*, Barcelona, Edicions 62.
- Vallverdú, Francesc (1990), *L’ús del català, un futur controvertit: Qüestions de normalització lingüística al llindar del segle XXI*, Barcelona, Edicions 62.
- Vargas Llosa, Mario, et al. (2008), *“Manifiesto por una lengua común”*, El País, http://elpais.com/elpais/2008/06/23/actualidad/1214209045_850215.html (last accessed: 7.11.2018).
- Vidal, Pau (2015), *El bilingüisme mata*, Barcelona, Pòrtic.
- Vila i Moreno, F. Xavier (2004a), *De l’ús al coneixement. Algunes reflexions sobre la promoció de la llengua al sistema educatiu. Jornades de la secció filològica de l’Institut d’Estudis Catalans a Vic*, Barcelona/Vic, Institut d’Estudis Catalans/Universitat de Vic/Ajuntament de Vic, 149–169.

- Vila i Moreno, F. Xavier (2004b), *Language, Education and Ideology in an Integrationist Society*, in: Els Witte et al. (edd.), *Contactforum: Language, Attitudes, & Education in Multilingual Societies*, Brussel, Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, 53–86.
- Vila i Ros, David (2007), *La llengua comuna*, Avui, 27.
- Viladot, M. Àngels (1993), *Identitat i vitalitat lingüística dels catalans*, Barcelona, Edicions Columna.
- Woolard, Kathryn A. (1989), *Double Talk: Bilingualism and the Politics of Ethnicity in Catalonia*, Stanford, Stanford University Press.
- Woolard, Kathryn A. (1998), *Introduction: Language Ideology as a Field of Inquiry*, in: Bambi B. Schieffelin/Kathryn A. Woolard/Paul V. Kroskrity (edd.), *Language Ideologies: Practice and Theory*, New York, Oxford University Press, 3–47.
- Woolard, Kathryn A. (2016), *Singular and Plural: Ideologies of Linguistic Authority in 21st Century Catalonia*, New York, Oxford University Press.
- Woolard, Kathryn A./Ribot Bencomo, Aida/Soler Carbonell, Josep (2014), *What's So Funny Now? The Strength of Weak Pronouns in Catalonia*, *Journal of Linguistic Anthropology* 23/3, 127–141.
- Wurl, Ursula M. (2011), *El concepte jurídic de llengua pròpia*, *Llengua i Dret* 56, 3–64.

Joan Pujolar

24 Migration in Catalonia: Language and Diversity in the Global Era

Abstract: In this chapter, I reflect on the debates on language and identity brought about by foreign migrants in Catalonia at the turn of the 21st century. I use a variety of information sources, including my own ethnographic study on language and immigration in Catalonia, as a particular case upon which to illustrate the terms of these debates in ways that I believe are relevant to most European linguistic minorities and mid-size nation-states or *majorités fragiles* (McAndrew 2010). I address two main aspects: a) how new discourses on integration and “social cohesion” were developed by different institutional actors in the early 2000s to reframe representations of national language and identity, b) how local communities incorporated these discourses in contradictory ways in their daily practices in their efforts to assert the primacy of local identity in the face of immigration. This will be based on an overview of language campaigns by the government of Catalonia from the 1980s to the present, on a commentary of the debate on Catalan as *llengua pròpia* or *llengua comuna*, and on data that I collected in 2002 in a language school for adult immigrants run by a church-based organization. My analysis will gradually move from the representations of immigration and language found in governmental language policies and linguistic activist organizations to the detail of the face-to-face social practices of language choice commonly seen in everyday life.

Keywords: Catalan, migration, multilingualism, nationalism, globalization

1 Introduction

In this chapter, I reflect on the debates on language and identity brought about by foreign migrants in Catalonia at the turn of the 21st century. The presence of populations that are visibly different and speak different languages has brought back the debates on the notion of “integration”. This term carries with it implicit ideas about how the national community is (re)imagined, and language often features prominently in this picture. Up to the mid-1990s, sociolinguistic debates had been restricted to the uses and spaces of Catalan and Spanish, and had not envisaged the possibility that speakers of many other languages might be present.

I use a variety of information sources, including my own ethnographic study on language and immigration in Catalonia, as cases upon which to illustrate tensions that I believe are relevant to most European linguistic minorities and mid-size nation-states or *majorités fragiles* (McAndrew 2010). In the case of Catalonia, there is the

<https://doi.org/10.1515/9783110450408-032>

added interest that immigration had been the object of much debate in earlier decades due to the important presence of workers from other areas of Spain. The migration flows of the late 1990s/early 2000s, often called “new immigrants”, present interesting differences in the ways that they have been portrayed, understood, represented and addressed in policy at various levels. In this chapter, in any case, I focus specifically on language: a) how new discourses on integration and “social cohesion” were developed by different institutional actors in the early 2000s to reframe representations of national language and identity, and b) how local communities have incorporated these discourses in contradictory ways in their daily linguistic practices as they seek to (re)assert the primacy of local identity and local understandings of nationality in the face of immigration.

The main materials examined will be a) the language campaigns carried out by the government of Catalonia from the 1980s to the present, b) a commentary of the debate on Catalan as *llengua pròpia* (‘the country’s own language’) or *llengua comuna* (‘common language’), and c) data that I collected in 2002 in a language school for adult immigrants set up by a church-based organization. My analysis will gradually move from the representations of immigration and language found in governmental language policies and linguistic activist organizations to the detail of the face-to-face social practices of language choice commonly seen in everyday life. First I will briefly address how the Catalan experience fits with contemporary debates on language and immigration in sociolinguistics. Then I will describe how political institutions and a significant sector of language activists have redefined the position of Catalan as an emblem of national identity in the context of increasing cultural diversity and multilingualism. Thirdly, I will describe how these shifts have been made operative in the NGO sector, as it contains organizations that now characteristically deliver an important part of the welfare services to immigrants in these neoliberal times. Afterwards, I will provide some glimpses of how the local population “treats” immigrants linguistically in everyday life. In the concluding section, I will try to analyze the continuities and contradictions that we find in how language is used and represented by different actors across all these different contexts.

2 Migration, nation state ideologies, and globalization

Since the term “globalization” established itself in the social sciences vocabulary and in the public imagination in the 1990s, there has been a significant shift in how issues of migration are (predominantly) addressed. The need to “integrate” or “assimilate” immigrants no longer constitutes the main logic behind the works of leading sociologists or anthropologists. Thus, Vertovec’s (1999) polyhedral notion of “transnationalism” has expressed an orientation to studying migration in terms of populations

subjected to “multiple ties and interactions linking people or institutions across the borders of nation-states” (1999, 447). Appadurai (1996) characterizes the contemporary world in terms of “flows” at many dimensions, i.e. products, capital, communication, and *people*. This new orientation has dominated sociolinguistic work too. The works of Zentella (1997), Blommaert/Verschueren (1998), and Heller (1999) marked a turn in which the emphasis was put on the ways in which language issues get mobilized by autochthonous groups to reproduce racial prejudice and social inequalities, i.e. to prevent effective political and economic participation by mobile populations. More recently, Blommaert/Rampton (2011) incorporated Vertovec’s notion of “superdiversity” to characterize the increasing diversity of contemporary populations in Europe criss-crossing cultural, racial, religious and economic boundaries. Garcia’s (2009) notion of “translanguaging” – the social practice of combining resources from different linguistic varieties in speech or writing – posits a further challenge to inherited ideas about linguistic boundaries that were inspired on the assumption that “nations” roughly represented the geographical expanse of relatively homogenous groups who naturally spoke separate languages.

These new concepts are intended to displace the assumption that traditional nation states should ideally be regarded as ethnically and culturally homogeneous. As such, we should treat the social sciences and sociolinguistics in particular not just as neutral analytical tools to understand the issue at hand; but also as an area of social practice that actively contributes – together with other actors – to define how political issues are addressed. From this viewpoint, academic perspectives provide part of the material that social actors use to build their understanding of the social world they live in. In the Catalan case, as we will see below, there has been an effort to make compatible multiculturalist perspectives with the promotion of Catalan as a public language in competition with Spanish. The term “multiculturalism” has circulated with very diverse meanings and uses and here I use it just to express the idea that the cultural capital and dispositions that immigrants display should no longer be treated simply as a problem, but as something that needs to have a space in current society. I intend to argue that immigration in Catalonia provides a case that can help to make the purview of the sociolinguistics of multiculturalism a little more complex and nuanced. Although the forces and interests that pursue the subjection and marginalization of immigrants are strong, we find in this part of the world – as in others – many social actors who seek to create political alliances and social complicities that provide specific spaces or opportunities of agency for immigrants. From this viewpoint, it is also easier to imagine and anticipate possibilities to foster social change.

It is worth recalling that narratives about the diversity of origins of national groups are not exceptional, as the well-known topos of the “melting pot” illustrates. Catalonia also has an important strand of thought that locates the nation’s pride as a “land of immigration”. The work of demographer Anna M. Cabré (1999) lent scientific weight to this view by representing the process of Catalonia’s modernization as marked by successive immigration “waves” since the 16th century, and particularly in

the 20-year periods around 1890, 1920 and 1960. Cabré claimed that it is thanks to immigration that Catalonia has been able to sustain rates of population and economic growth similar to other industrial nations. She also suggested that immigration is deeply inscribed in Catalan culture because the strategies of wealth preservation by families encouraged intermarriage with outsiders. The title of Pernau's (1995) collection of short autobiographies by (non-Catalan born) political and cultural leaders is in this sense expressive of this view of the national community as made up of immigrants: *el somni català* 'the Catalan dream'.

So far, however, immigrants had primarily come from other regions of Spain or from the South of France, and they shared with Catalans key cultural features such as race or religion (Pujolar 1995), and very closely-related languages (Catalan-speaking immigrants from Valencia or the Balearic Islands have always passed fairly unnoticed, particularly as language became the main index of differentiation). During the 20th Century, the term *català* 'Catalan' became a way to name the locals, i.e. those who spoke the language, as opposed to outsiders, who were called *castellans*, i. e. all those who spoke Castilian whatever their origin. Derogatory terms such as *murcianans* or *xarnegos* were semantically similar, although more charged with class prejudices.

The latest migration episode around the year 2000 has been different in that most migrants came from other continents, thus presenting differences not just of language but also of race and religion. According to official figures, 14 % out of 7,5 M. were registered in 2016 as having another nationality: 21 % of them from Latin American countries, 21 % from Morocco, 9 % from Romania, 5 % from China, 5 % from Italy and 4 % from Pakistan. An additional 18 % of the population of Catalonia has been born in other Spanish regions (source: www.idescat.cat). This means that, were Catalonia an independent country, its foreign population (32 %) would lead the ranking of OECD countries just after Luxembourg (44 %, source: OECD 2016). Undocumented migrants and the large numbers of naturalized people from Latin America could even increase these rates significantly.

Thus the slogan that depicts Catalonia as a "land of immigrants" has actually a strong material basis. There are however telling contrasts in how migration has been historically constructed across by the Catalan elites. Up until the 1930s, discourses on immigration were largely portraying the phenomenon as a potential threat to national identity. This changed significantly during the 1960s and 1970s as Catalan nationalists and left-wing activists built alliances with immigrants in the opposition movements against dictator Franco. In Pujolar (1995) I provided an overview of already rich and long debates on migration that have not ceased within the fields of sociology, anthropology and sociolinguistics in Catalonia. However, as Fernández (2008) has shown in his appraisal of 25 years of the Catalan cultural field, the theme of migration is surprisingly subdued in cultural production and literature. Fernández suggests that the *erasure* of the immigrant experience in the arts also makes a constitutive feature of Catalan society and culture, which reveals contradictions with the narratives of Catalonia as a "land of immigrants".

3 From national identity to social cohesion

One important development with regard to immigration issues is how the Generalitat de Catalunya (the regional autonomous government), together with language activist groups, has reconstituted earlier discourses about the Catalan language during the last two decades. From an original classical approach to language as an emblem of national identity, in the face of increasing transnational migration, language policy actors gradually elaborated the idea that language was a ‘means of social cohesion’ and that Catalan had to be presented as the land’s ‘common language’ rather than the established term *llengua pròpia*, ‘(the country’s) own language’.

I have described elsewhere (Pujolar 2010) how the governmental discourses that articulated language, culture and national identity gradually changed from a classical nationalist frame that was hegemonic during the 1970s and 1980s to a more nuanced narrative in which Catalan language and culture are now presented as a meeting point in a scenario characterized by diversity. Through the centrality of Catalan, I argued, institutions could deliver a public good labeled ‘social cohesion’. It is useful in this regard to compare governmental white papers such as the *Pla de Normalització Lingüística*, the ‘linguistic normalization plan’ (Departament de Cultura 1992) and its projection in the Department of Social Welfare (Departament de Benestar Social 1992), to the productions in the ensuing decade: the plans for ‘language and social cohesion’ (Departament d’Educació 2004; 2007). In the earlier documents there was not a single reference to migration and the government was primarily concerned in getting its own civil servants to use Catalan officially. The goal was ‘to promote the use of Catalan, as Catalonia’s own language, in the main sectors of society’ (Departament de Cultura 1992, 6). Tellingly, the latter documents were produced by the Department of Education rather than the more peripheral Department of Culture. One important aspect of the plan was to abandon an earlier “compensatory” model in which newly arrived migrants were segregated into special classrooms. During the early 2000s the new model of the *Aules d’acollida* ‘reception classrooms’ turned these spaces into resources used only exceptionally and resources were devoted to support the students’ incorporation into the mainstream classrooms (Pujolar 2010).

The shift was also noticeable in the reorientation of the institutional publicity campaigns. The 1982 initial slogan *El català, cosa de tots* ‘Catalan, a thing for all’ and the 1985 *Depèn de vostè* ‘it is up to you’ were intended to establish the linguistic cohabitation of local Catalan and Spanish speakers. Catalan speakers were called to be more active users of the language, and Spanish speakers were called to own it too as part of the nation. The term “immigrant” was never used given that hegemonic discourses portrayed Catalans as *Un sol poble*, ‘a single people’ (and Spanish-speakers were Spanish nationals anyway). In contrast, the 2003 slogan *Tu ets mestre* ‘you are (a) teacher’ was accompanied with posters and videos in which local and foreign characters were clearly set apart with different roles. In 2005, *Dóna corda al català* ‘give Catalan a wind-up’ addressed language learners specifically. It featured a toy, a

walking mechanical denture that symbolized the experience of learning and using a new language that needed constant effort. In 2009, *Encomana el català* ‘make Catalan contagious’ also concentrated its attention on foreign migrants and on the need to get local Catalan speakers to support learning.

In addition to governmental initiatives, similar shifts occurred within the complex web of Catalan political and cultural activism. One of the key debates has been around the concepts of *llengua pròpia*, an expression designating the territory’s ‘own language’, and that of *llengua comuna* ‘the common language’. Riera Gil (2013) provides a good appraisal of debates amongst academics – largely between law specialists and political scientists. Jou (2012) argued that the territorial dimension implicit in the concept of *pròpia* allowed to develop policies advantageous to the use of Catalan, whereas Branchadell (1997) criticized the concept as ‘illiberal’. Grassroots activists have shown little interest for these debates in themselves; but they have overwhelmingly turned to the concept of *llengua comuna*, a notion initially developed within a group called *Veu pròpia* ‘a voice of our own’ that gathered ‘new speakers of Catalan’, i. e. speakers who had other native languages (see *Plataforma per la Llengua* 2010). The *Plataforma per la llengua*, a Barcelona-based language activist group, gathered the support of 24 NGO representing different immigrant communities in 2008 for the manifesto *El català, llengua comuna* (*Plataforma per la Llengua* 2008) in which newcomers ‘offer our languages and cultures of origin to enrich this country’, affirm that ‘Catalan is the language that unites us all’ and that ‘defending Catalan is defending all our languages’. Here the use of a first person ‘we’ and ‘us’ that denotes migrants constitutes a significant shift in the discourses on the Catalan language. As I had argued elsewhere (Pujolar 2007b), language debates had been predominantly formulated from the perspective of an ‘us’ that pointed exclusively to native speakers. It is also noticeable that the cultural capital of immigrants was presented as an asset that could be ‘offered’, thus adopting a multiculturalist standpoint. The manifesto for the ‘common language’ also characterized Catalan as a ‘language of social inclusion’ and as ‘synonymous of cohesion’ in a text which also used the phrase *llengua pròpia* a few times but never the concept of “integration”. The slogan *El català llengua comuna* has been repeatedly hailed as the motto for the *Plataforma*’s celebration of Saint George’s day, the de facto day of national culture (<https://www.plataforma-llengua.cat/sant-jordi>), whether the activities have involved demonstrations, flash mobs, lip dubs, mural mobs, etc...

Attention to immigration and to linguistic diversity has however a weak spot when it comes to providing explicit accounts on the position of Spanish or Castilian in Catalan society. In official documents, references to Spanish are few and noncommittal. This reflects, I would argue, an underlying tension in that the celebrations of linguistic diversity are primarily invested in securing the position of Catalan in public life before increasing evidence that newcomers learn Spanish first as a *lingua franca* (Fukuda 2017; Lapresta-Rey/Hugué/Fernández-Costales 2016; Sabaté i Dalmau 2016), and bearing in mind that current Spanish regulations require passing a Spanish A2

level exam to obtain the nationality. Sabaté i Dalmau (2009) provides a good account of how these tensions play out in Catalan schools, in which the willingness to legitimize and visibilize immigrant languages must be modalized when it comes to Castilian.

4 Language, citizenship, and welfare

The new language policy paradigm has been projected onto the operating ways of particular institutions and bureaucratic procedures. It has affected in particular the adult language education provided by the public and the private (NGO) sectors. Of particular interest are the ways in which language teaching gets mobilized as the means to grant access to welfare services and to ensure that immigrants are socialized into the specific forms of self-management and entrepreneurship that have developed in neoliberal societies: birth control, health monitoring, nutrition, job seeking, training, and so on. Additionally, there is the issue of how language is inserted in the procedures to obtain nationality and also residence and work permits. From this viewpoint, Spain has followed the European trend to set a language exam to obtain nationality (Ley 19/2015 de medidas de reforma administrativa, 2015, *Disposición final séptima. Procedimiento para la obtención de la nacionalidad española por residencia*). The Generalitat also mentions knowledge of the two languages as an aspect to bear in mind in official reports evaluating the ‘degree of social integration’, which then are used by the Spanish government to grant work and residence permits. This acts as an incentive to immigrants to take Catalan classes although the criteria finally used by officials to evaluate linguistic competence are rather flexible, not to say vague (Direcció General per a la Immigració 2014).¹

Groups of linguistic activists and the platforms created by migrants connected to specific nationalities are particular examples of a much wider and diverse network of actors that participate in immigration policies. As both Spanish and Catalan institutions took up the neoliberal turn at the end of the 20th century, they increasingly relied on many types of NGOs to provide specific services connected to welfare such as legal advice, housing or food relief (see Codó/Garrido Sardà 2010; Garrido Sardà 2014; Pujolar 2007a). Trade unions, local catholic branches of Caritas Diocesana, the Red Cross, the transnational organization Emmaus, or more local groups, they all became major actors in the implementation of migrant settlement policies, and many of them engaged in the organization of language classes for adults. Although the Catalan government had a sizable network of ‘normalization centers’ that offered free language courses, this official consortium could barely cope with the peak of 127,807

¹ I am thankful to Mr. Orland Cardona for his help in clarifying current legal and administrative procedures.

registered students in the 2009–2010 course year. As the migration flows slowed down and spending cuts arrived, the figure was brought down to 76,939 in 2017 (Consorci per a la Normalització Lingüística 2018). In any case, the official language schools and local adult education services never covered the demand either. As a result, non-governmental settlement services often engaged in Catalan language tuition particularly addressed to the most vulnerable groups, i.e. people with visa problems or recently arrived women who stayed at home on their own most of the time minding small children (Pujolar 2007a).

This situation provided new settings in which the discourses over Catalan were reconstituted by different actors who had different objectives in which promoting the language as such was not seen as central. Indeed, the first independent charities that had begun providing language courses to newly arrived migrants in the early 1990s had offered mainly Spanish classes (Rovira et al. 2004). The main argument for this choice was that Spanish was generally more instrumental to accomplish the basic communicative needs of immigrants at the workplace and with the administration (Pujolar 2007a; 2010). It was certainly true that newcomers had easier access to Spanish for many reasons. At first most came from Morocco, whose Northern population traditionally speaks some basic Spanish. Many of them had spent months working in other parts of Spain before their arrival to Catalonia and were planning to still move on to France or further. Most had not been aware that the Catalan language even existed before their arrival, and some did not always notice at first the different languages spoken by the local population (Caglitütüncil Martínez 2016). Additionally, their administrative dealings with residence and work permits were still handled by the Spanish administration. Catalan civil servants (e.g. municipal workers, health workers) would not require them to use Catalan either. In fact, most of the local population (neighbors, shopkeepers, employers), even in places where Spanish was hardly heard, would address them in Spanish as was customary to anyone perceived as “foreign”. In this context, social workers and NGO leaders had developed an argument that Spanish was the language that people “really” needed for practical purposes, whereas the value of Catalan was merely “symbolic”, which in common parlance is a synonym of “ornamental” or “non-functional”.

The sociolinguistic quandary was however more complex than that, as it affected the “symbolic” aspects of language in ways that were materially consequential. To put the question simply: what are the chances of social and economic success (in Catalan-speaking contexts) for foreign people who have learned just Spanish and not that fluently? In the Northern town of Salt in 2001, for instance, a conflict emerged in a public meeting in which some immigrant groups demanded that Spanish be used instead of Catalan. The event triggered a protracted national debate in the newspapers that I have reported elsewhere (Pujolar 2010), and it effectively showed that the supposed functionality of Spanish also had limits when it came to participating in public life. There were others: migrant children were taught in Catalan and their parents found it more necessary to master it to attend school meetings or help their

children with the homework. And those who learned Catalan suddenly found that local people were noticeably more communicative or even more willing to rent their apartments or to trust them with better jobs.

In response to these debates and to the articulation of the official discourse on social cohesion, many of the NGOs effectively appropriated the view that learning Catalan made integration easier and more effective. In the first decade of the 2000s many Spanish courses were substituted by Catalan courses (Codó/Garrido Sardà 2010; Garrido Sardà 2010).

However, this did not fully solve the contradictions that existed when it came to providing immigrants with the means to participate in Catalan society, be these through employment, community life, public life, politics or as users of social services. With rare exceptions, services to immigrants were still provided in Spanish whether this was in the public health system (Moyer 2011), in trade unions (Codó/Garrido Sardà 2010) or in the other NGOs mentioned above. As Garrido Sardà (2010) aptly depicts, learning Catalan and using some Catalan expressions was presented as means to display a 'willingness to integrate' before local audiences; but otherwise Spanish was retained as the means to conduct everyday communication and get things done.

In the catholic NGO where I conducted an ethnography in the early 2000s in a Catalan-speaking town, I could ascertain that language classes were provided in Catalan; but the very same students who took Catalan classes were addressed in Spanish by everyone except the teachers: the childminders who took care of their children during class, the police officers who came in to make presentations about their services, diet or health professionals who delivered advice about household habits, etc. It was interesting to see that the language classes were delivered in Catalan, as it seems, as an expression that these institutions were facilitating some form of "integration"; but that Spanish was still used for everything that had to do with participation with state and welfare institutions that would supposedly make immigrants effective citizens in practical or ordinary terms. And all those who spoke Spanish to immigrants were people who otherwise interacted in Catalan amongst themselves.

These unresolved tensions thus expressed the ambivalence that immigrants met when it came to working out the language or languages through which they could access citizenship both from the perspective of official procedures (work and residence permits and visas) but also as social actors who must engage in the multiple aspects of public life (participate in decisions, sign contracts, access services) in formal and informal contexts. In theory, both Catalan and Spanish worked. In practice, neither fully did on its own.

5 The politics of language and subjectivity

The political consensus that Catalan was important for ‘integration’ or ‘cohesion’ was generally contradicted not in explicit ways or in formal podiums; but through the spontaneous practices of language choice by most Catalan-speakers in everyday life. Through the ordinary micropolitics of social interaction, language becomes a salient resource to reproduce or challenge inherited notions of what it is to be a local/national and to be a foreigner. In the Catalan-speaking territories, the value of the language as a local resource often constitutes an obstacle to its use as a lingua franca; but it can also be used to index different possibilities or degrees of belonging, as when people distinguish between those who are recent arrivals or whose status is provisional from those who are seen as ‘settled’ and hence committed to the place and its ongoing struggles. Thus, speaking Catalan or Spanish to ‘immigrants’ becomes a highly meaningful choice connected with the subtle everyday-life gestures of inclusion and exclusion.

Local, Catalan-speaking, people were therefore organizing their language choices in divergent contradictory ways. Old habits inherited after years of political repression led most people to treat Spanish as a more accessible language. This is because besides the explicit principles fostered by institutions and nation-wide NGOs, there is a domain of politics that is arguably inscribed in people’s bodies and which rules their perceptions, interpretations and reactions at a subtle phenomenological level. Despite all the discourses of meeting points, common languages and intercultural understanding, most people will only speak Catalan to those whom they consider to be Catalan, and this does not include people who show heavy accents, unexpected clothing, or dark-colored skins. And just like the old divide between ‘Catalans’ and ‘Castilians’ has always lingered in people’s consciousness, with new immigrants the Spanish language is back for many to express foreignness.

This is not to say that these ‘habits’, as they are popularly called, cannot be changed. It was teachers, some social servants and (ironically) nationalists who trained themselves to abandon this convention and speak Catalan. In an adult school close to the site I researched, I found that the staff had taught themselves to respond consistently in Catalan and to speak more clearly and slowly (the special register of how to talk to foreigners in Catalan is generally absent from most people’s repertoires). I have also seen the workers of the normalization consortium practice a fine discipline in their language choices.

In the specific NGO that I researched, however, most people were volunteers and there was a limit to how the organization could intervene in people’s behavior. So, teachers of Catalan had made the effort to readjust their linguistic habits; but only when they talked to the specific students who attended their class. We can see this in the example below, when the elderly grandmother of one of the students entered the premises:

An episode: The visit of the grandmother

The class was practicing spelling in a routine way when Fadma, a young Moroccan student, entered accompanied by her grandmother. Mara, Alija and Racquel were other Moroccan students already in the classroom.²

- 1 Carme: aha aixís · · hi posarem · · ensenya-li · Cris · una ratlleta
- 2 ‘Aha. That way. We shall write... Show her, Cris: one line...’
- 3 Fadma: holaa [ambilingüe, obrint la porta de l’entrada i entrant]
- 4 ‘Hello [ambilingual, opens the outside door and comes in]’
- 5 Carme: hola · mira [ambilingüe]
- 6 ‘Hey, look! [ambilingual]’
- 7 Fadma: >(qué) [Manté la porta oberta mentre una dona marroquina molt gran entra vestida amb xil·laba i ranquejant]
- 8 ‘(What!) [Keeps the door open while an old Moroccan woman comes in behind her, dressed with a djellaba] and limping heavily’
- 9 Carme: · <Fadma · bon dia · [en castellà] *hombree* ·
- 10 ‘Fadma! Good morning. [in Spanish] *Maaan!*’
- 11 Fadma: >*mi abuela* (xxxxx)
- 12 ‘>*my grandmother* (xxxxx)’
- 13 Racquel: [rient] <*la abuela también*.
- 14 ‘[laughing] <*the grandmother too*.’
- 15 Algú: <holaa [ambilingüe, la gent riu]
- 16 ‘<Hello [ambilingual, people laugh]’
- 17 Carme: *muy bien · venga* · ·
- 18 ‘*Very good. Come on* · ·’
- 19 Mara: [s’exclama en àrab] haad-i ki-daar-t daar-t[?] > · · (xx) Aa[!]
- 20 ‘[in Arabic to herself] How come did this woman manage to > (xx) · Ah!’
- 21 Carme: <*venga* ui que bé · posem cadires [rialles] · · mira *la abuela* molt bé
- 22 [Irene s’aixeca, dona la mà a la dona i acota el cap; Racquel fa el mateix després sense aixecar-se]
- 23 ‘<*Come on*. This is really nice. Let’s get some chairs [laughters]. · Look *the grandmother*. Very good!
- 24 [Irene, the research assistant, stands up and shakes the woman’s hand; Racquel also shakes her hand without standing up].’

In lines 9, 17 and 21 Carme – the teacher – uttered in Spanish the greetings that were supposed to be produced for the benefit of the grandmother. Not that the grand-

² Catalan is in normal type, Spanish in italics, Moroccan Arabic underlined and the English translation is enclosed in single quotation marks. The > signs point to the fragments of talk that overlap with those beginning with <. Raised dots “·” indicate pauses of half a second.

mother had actually given any signs of speaking or even understanding Spanish. In fact, in the little while during which she lingered, she did not utter a word in any language, neither did she make eye contact with anyone or expressed that she was listening. She just sat and stood silently there and was treated as a welcome visitor although neither she nor her granddaughter produced any explanation or indication as to why she had come.

Clearly, a rational appraisal of what language was best to use in the situation was not done. And it was not done either by the policemen or the social workers who led some of the sessions. Interestingly enough, they often asked in these occasions which language was best to use, as a gesture to express that Catalan was not fully ruled out. But they normally answered the question themselves without allowing much discussion and went on in Spanish. Had they really checked, as I did, they would have discovered that the audience was often composed by students of Catalan with a few additions whose proficiency in Spanish was just as weak as in Catalan.

The logics of belonging and integration that were thus expressed did not affect only language choice, but other subtle aspects in which a certain colonial ethos was reproduced. At one point, during a language exercise, one student said that she gave her toddler milk to drink when he had a fever; and the teachers suddenly broke the pedagogical frame to impart advice about the best childcaring practices in the situation. Although the mother was comparably younger and such situations are possible amongst locals too, the tone and the frequency of these occasions expressed the idea that locals were not just teachers of language but also of modern liberal citizenship in which individuals must gain autonomy with the help of local institutions and the assistance of the locals: e.g. in a discussion session amongst the NGO's volunteers, significant time was devoted to the issue of training students to manage punctuality, order in the classroom, and regular attendance.

6 Conclusion

The Catalan case provides one case study to understand how globalization is putting traditional ideologies of language and identity under pressure. Although Catalonia has historically been a traditional receptor of migrants, the last significant influx is being experienced as markedly different because people come from places more distant both geographically and culturally, and because the immigrant population is much more diverse internally. Moreover, their patterns of settlement are not comparable as migrants nowadays display trajectories of ongoing mobility. The intellectual context has also changed and there are attempts to resignify the national community in ways that provide room for cultural difference, as the discourses over 'social cohesion' and *llengua comuna* show in some aspects.

However, one important aspect of this change does not rely so much on the characteristics of the new mobile citizens, but on the way in which the state has

restructured its operations and externalized its policies to actors which may share its institutional political agendas only partially. Not all NGOs have issues of language and identity as their priority so that they can choose to enact official policies in multiple ways that can potentially contradict their explicit objectives, in this case, to promote Catalan as the 'common' language of social cohesion.

Finally, the fact that Spanish is much more widely used as a lingua franca amongst locals and immigrants and amongst immigrants of different origins constitutes one more element of contradiction between national wishes and practical realities. The habit or instinct that locals have to speak Spanish to whomever they perceive as being 'non-Catalan' helps reproduce this situation; but it is not unlike what happens in other European contexts in which languages such as English, French or Russian also gain space as lingua francas. From this viewpoint, what we see is the emergence of specific linguistic economies in context of mobility that transcend the former logics of national languages. It remains to be seen how these phenomena develop in the future, particularly when subsequent generations will eventually have experienced (as they do in Catalonia) the linguistic market still enforced by national school systems.

7 Bibliography

- Appadurai, Arjun (1996), *Modernity at Large: Cultural Dimensions of Globalization*, Minneapolis, University of Minnesota Press.
- Blommaert, Jan/Rampton, Ben (2011), *Language and Superdiversity*, *Diversities* 13/2, 1–22.
- Blommaert, Jan/Verschueren, Jeff (1998), *Debating Diversity: Analysing the Discourse of Tolerance*, London, Routledge.
- Branchadell, Albert (1997), *Liberalisme i normalització lingüística*, Barcelona, Empúries.
- Cabré, Anna M. (1999), *El sistema català de reproducció: cent anys de singularitat demogràfica*, Barcelona, Proa.
- Caglitütüncgil Martínez, Tülay (2016), *Lessons from the Linguistic Agencies of Immigrant Women: A Critical Ethnography of Education at the Social Integration Programs in Madrid and Barcelona*, Ph. D. Thesis, Barcelona, Universitat Oberta de Catalunya.
- Codó, Eva/Garrido Sardà, M. Rosa (2010), *Ideologies and Practices of Multilingualism in Bureaucratic and Legal Advice Encounters*, *Sociolinguistic Studies* 4/2, 297–332.
- Consorci per a la Normalització Lingüística (2018), *Memòria 2015 – resum*, Barcelona, [https://arxius.cpnl.cat/arxius/memories/Memoria 2017- resum.pdf](https://arxius.cpnl.cat/arxius/memories/Memoria%202017-resum.pdf) (last accessed 03.12.2018).
- Departament de Benestar Social (1992), *Pla de normalització lingüística del Departament de Benestar Social 1992*, unpublished report, Barcelona.
- Departament de Cultura (1992), *Pla de normalització lingüística: 1992*, unpublished report, Barcelona.
- Departament d'Educació (2004), *Pla per a la llengua i la cohesió social*, unpublished report, Barcelona.
- Departament d'Educació (2007), *Pla per a la llengua i la cohesió social. Educació i immigració*, unpublished report, Barcelona.
- Direcció General per a la Immigració (2014), *Instrucció DGI/BSF/1/2014 per la qual s'estableixen els criteris generals per a l'elaboració dels informes d'estrangeria competència de la Generalitat de Catalunya*, Barcelona, Generalitat de Catalunya.

- Fernández, Josep-Anton (2008), *El malestar en la cultura catalana: la cultura de la normalització*, Barcelona, Empúries.
- Fukuda, Makiko (2017), *Double Gateway to the Host Society? Knowledge and Perceptions of Japanese People Living in Catalonia Regarding Language*, *Journal of Multilingual and Multicultural Development* 38/1, 19–34. DOI 10.1080/01434632.2016.1146288.
- García, Ofelia (2009), *Education, Multilingualism and Translanguaging in the 21st Century*, in: Tove Skutnabb-Kangas (ed.), *Social Justice Through Multilingual Education*, Clevedon, Multilingual Matters, 140–158.
- Garrido Sardà, M. Rosa (2010), *Tensions ideològiques sobre “la llengua d’acollida” en un projecte residencial per a migrants*, *LSC– Llengua, Societat i Comunicació* 8, 19–26.
- Garrido Sardà, M. Rosa (2014), *Emmaus as a Transnational Imagined Community Language, Interdiscursivity and Stratification in a Social Movement*, Ph. D. thesis, Universitat Autònoma de Barcelona, <https://www.tdx.cat/handle/10803/285359> (last accessed 18.03.2019).
- Heller, Monica (1999), *Linguistic Minorities and Modernity: A Sociolinguistic Ethnography*, London, Longman.
- Jou, Lluís (2012), *Conceptes jurídics i pràctica política. Les conclusions de l’Àrea de Llengua del II Congrés internacional de llengua catalana vint-i-cinc anys després*, *Revista de Llengua i Dret/ Journal of Language and Law* 58, 53–73.
- Lapresta-Rey, Cecilio/Huguet, Àngel/Fernández-Costales, Alberto (2016), *Language Attitudes, Family Language and Generational Cohort in Catalonia: New Contributions From a Multivariate Analysis*, *Language and Intercultural Communication* 17/2, 1–15. DOI 10.1080/14708477.2016.1165241.
- Ley 19/2015 de medidas de reforma administrativa (2015), *Ley 19/2015, de 13 de julio, de medidas de reforma administrativa en el ámbito de la Administración de Justicia y del Registro Civil*, Madrid, Boletín Oficial del Estado, 58125–58149.
- McAndrew, Marie (2010), *Les majorités fragiles et l’éducation: Belgique, Catalogne, Irlande du Nord, Québec*, Montréal, Les Presses de l’Université de Montréal.
- Moyer, Melissa G. (2011), *What Multilingualism? Agency and Unintended Consequences of Multilingual Practices in a Barcelona Health Clinic*, *Journal of Pragmatics* 43/5, 1209–1221.
- OECD (2016), *International Migration Outlook 2016*, Paris, OECD Publishing. DOI 10.1787/migr_outlook-2016-en
- Pernau, Gabriel (1995), *El somni català*, Barcelona, La Campana.
- Plataforma per la Llengua, La (2008), *Manifest “El català, llengua comuna”*. Sant Jordi 2008, Barcelona, Plataforma per la Llengua.
- Plataforma per la Llengua, La (2010), *El català també és meu*, Barcelona, Fundació Vincle.
- Pujolar, Joan (1995), *Immigration in Catalonia: The Politics of Sociolinguistic Research*, *Catalan Review*, *International Journal of Catalan Culture* 9/2, 141–162.
- Pujolar, Joan (2007a), *African Women in Catalan Language Courses: Struggles over Class, Gender and Ethnicity in Advanced Liberalism*, in: Bonnie McElhinny (ed.), *Words, Worlds, and Material Girls: Language, Gender, Globalization*, Berlin, Mouton de Gruyter, 305–347.
- Pujolar, Joan (2007b), *The Future of Catalan: Language Endangerment and Nationalist Discourses in Catalonia*, in: Alexandre Duchêne/Monica Heller (edd.), *Discourses of Endangerment: Interest and Ideology in the Defense of Languages*, London, Continuum international, 121–148.
- Pujolar, Joan (2010), *Immigration and Language Education in Catalonia: Between National and Social Agendas*, *Linguistics and Education* 21, 229–243.
- Riera Gil, Elvira (2013), *Sobre el concepte polític de llengua comuna: una aproximació teòrica i comparada*, *Revista de Llengua i Dret/ Journal of Language and Law* 60/1, 91–110. DOI 10.2436/20.8030.02.31.

- Rovira, Marta, et al. (2004), *El català i la immigració. Anàlisi de l'oferta de cursos de català als immigrants adults extracomunitaris*, Barcelona, Mediterrània.
- Sabaté i Dalmau, Maria (2009), *Ideologies on Multilingual Practices at a Rural Catalan School*, *Sociolinguistic Studies* 3/1, 37–60. DOI 10.1558/sols.v3i1.37.
- Sabaté i Dalmau, Maria (2016), *El català dels locutors: pràctiques i ideologies lingüístiques de la migració en espais urbans castellanoparlants*, *Treballs de Sociolingüística Catalana* 26, 105–121. DOI 10.2436/20.2504.01.111.
- Vertovec, Steven (1999), *Conceiving and Researching Transnationalism*, *Ethnic and Racial Studies* 22/2, 447–462. DOI 10.1080/014198799329558.
- Zentella, Ana-Celia (1997), *Growing Up Bilingual: Puerto Rican Children in New York*, Oxford, Blackwell.

25 Catalan Worldwide

Abstract: This chapter deals with the presence of the Catalan language in the world beyond the Catalan Countries, both historically through settlement, emigration and exile and in the present. The chapter highlights the organisation of Catalan communities in *casals* mainly in Europe and in America, but also on other continents, the presence of the language in the media and on the Internet, and the international network of Catalan studies and their publications.

Keywords: Catalan language, Catalan Casals, Catalan publications, Catalan media, Catalan studies

1 Europe

Beyond the Catalan Countries, the territories where the Catalan language is spoken, Catalan was a language of communication with considerable circulation in certain periods and regions (Coromines 1954, 35–52; Vallverdú 2001, 327–369). In addition, from the 20th century onwards, the activities of associations devoted to the study of the Catalan language and culture, included at present in the FIAC (Federació Internacional d'Associacions de Catalanística), and the network of Catalan lectorships at foreign universities, today coordinated by the Institut Ramon Llull, have ensured and amplified the presence of the Catalan language in the international academic world (Bover 1993).

During the Middle Ages and until the period of the Renaissance, Catalan remained a constant presence along the coasts of the Mediterranean and Southern Europe. Thanks to its active trade with the Eastern Mediterranean, Catalonia had an extensive network of funduqs or corn exchanges (*alfòndecs*) and consulates at its disposal. The oldest existing documentation of one of these outposts dates back to 1187 in the ancient Phoenician city of Tyre, and many of them remained in operation for centuries, such as the one in Alexandria in Egypt, which was still operating as late as 1498.

As early as the reign of James I, an expedition of Catalan crusaders was organised to conquer the Holy Land; the fleet was destroyed by a storm, but eleven ships reached Saint John of Acre. Both the great philosopher Ramon Llull and the chronicler Ramon Muntaner used the Catalan language constantly in their numerous voyages across the Mediterranean. Ramon Muntaner (1265–1336), for example, describes how the Italian personalities in the service of the Crown of Aragon, such as the great admiral Roger of Lauria or the rationalist teacher Corral Llança, had adopted Catalan

<https://doi.org/10.1515/9783110450408-033>

as a day-to-day language. In addition, in the Italic territories belonging to the Catalan-Aragonese Confederation, especially in the cities, traders and educated people knew and used the Catalan language. This Catalanisation was especially intense on the island of Sardinia (Carbonell 1984; Bover 2007, 1–73), where the Catalan language is still alive today in the city of l'Alguer/Alghero (Bover 2007, 75–235). The influence of Catalan originally extended over the whole island, and we have extensive evidence of it in old records, place names, family names, and the influence of Catalan on the Sardinian language, especially in the Campidanese dialect in the south of the island. During the 14th century, when Athens and the Morea were in the possession of Catalonia, the prestige of the Catalan language led to its use spreading all around the Aegean Sea. From the 15th century there are numerous eye-witness accounts of the popular enthusiasm with which the preachings in Catalan of Saint Vincent Ferrer were met, which were both understood and acclaimed in France and in Italy and, more surprisingly, also in England and Brittany. The Catalan language was also present orally and in documentation in the Kingdom of Naples during the rule of the Catalan-Aragonese, when citizens from the Catalan territories began settling there after King Alfons the Magnanimous was crowned King of Sicily in 1442. On the military side, special mention must be made of Catalonia's participation in the defence of Constantinople (1453) – of which we have several eye-witness accounts – this being a city where Catalan commerce had been important since the 13th century, in the period of the emperor Michael VIII Palaiologos, and where the presence of a Catalan consulate is recorded as early as 1268. In Rome, during the pontificates of the Valencian popes of the Borgia family, Calixtus III (1456–1458) and Alexander VI (1492–1503), people learnt, spoke and wrote Catalan, the mother tongue of the popes. The expansion and prestige of Catalan throughout the Mediterranean has left its mark through linguistic loans in many languages; Catalanisms are especially prominent in maritime vocabulary. The diaspora of the Catalan Jews abandoned the language after their expulsion when mixing with the wider Sephardi community, as did the Portuguese Jews, but they kept the national denomination and some variants of their language. In addition, at the end of the century, a considerable number of German and French printers learnt the Catalan language in order to establish themselves in several cities in the Catalan territories, and introduce printing there. The prestige of the language and its literature during the Middle Ages and the Renaissance, but also later on, has left an impression on several literatures. Catalan was a widely known language among the great Aragonese writers and early Spaniards, such as Juan Manuel, the Arcipreste de Talavera or the Marquis of Santillana, and some even wrote in Catalan, such as Enrique de Villena or Bartolomé de Torres Naharro. The works of Catalan authors were very quickly translated into other languages and we can see their influence on authors like Montaigne, Cervantes, Giordano Bruno, Leibniz, Bacon or the great Scottish philosophers.

After the War of Spanish Succession, the Catalan Austriacist exiles (1714–1715) had various destinations but mostly established themselves in Vienna, in Italy, especially

in Naples and Milan, in addition to Rome, where almost the entire ecclesiastical community was concentrated, and in Hungary, where troops on active service predominated. Between 1735 and 1738 hundreds of exiled Catalans were transported to the Banat of Temeswar, where they created the colony known as New Barcelona, which was never consolidated. Also during this century, on the island of Minorca – incorporated into the British Crown (Crespo i Sala 2018) – many members of the British garrison learnt and used Catalan.

Mention must also be made of the more recent settlement of Catalan fishermen along the estuaries of Galicia, where they set up the cannery industry, or of the fishermen from Sitges established in La Figuereta – today known as Isla Cristina – in Andalusia. Today Catalan gypsies have extended the use of the language in communities settled in several towns of Provence and other areas of France – recent studies reveal that, for instance, there are 4,000 regular Catalan speakers within the gypsy community of Montpellier alone (Escudero 2014; Casanova 2014, 2016). Similarly, exile and emigration, from the 19th century until the present day, and often organised in *casals*, has helped to spread the spoken use of Catalan around the world.

The scientific and academic study of the Catalan language and its literature throughout the 20th century in particular developed initially in Europe, and has given rise to the current diffusion of international Catalan studies.

2 America

Today there is widely held and well-grounded suspicion that Christopher Columbus was of Catalan origin, and indeed he was surrounded by a remarkable number of Catalans on his American expeditions. In spite of the serious limitations to trade with these territories that the citizens of the Crown of Aragon suffered, and which limited Catalan migration considerably until the 18th century, this continent boasts nearly a thousand Catalan place names, such as the island of Montserrat, one of the Lesser Antilles in the Caribbean, so christened in 1493. A large number of Catalan and Majorcan people took part in the colonisation of California during the 18th century. This process was headed by the governor Gaspar de Portolà, Pere Fages and his company of Volunteers of Catalonia, and Juníper Serra, religious head of the expedition. On the other North American coast, a group of Minorcan colonists established themselves near the bay of Mosquitos, in the then British province of East Florida in 1768. This colony of New Smyrna failed, and in 1777 the survivors took refuge in the ancient city of Saint Augustin, where some Minorcan surnames remain, albeit more or less transformed by the English language, along with a few place names, and where some trace of the Mahonese Catalan spoken by the old colonists can still be found in the local language, in addition to the *formatjades* and other culinary dishes (Rasico 1990). Throughout the 19th century, the island of Cuba was the main destination of Catalan emigration. It is estimated that one in every five migrants coming from Spain

was Catalan, and one in ten were soldiers. More than a quarter of the Havana city traders and well more than half of those in the Santiago area were also Catalan; in the island's Spanish, for instance, the word *catalán* took the meaning of 'shopkeeper'. In 1840, in the city of Havana, the Societat de Beneficència dels Naturals de Catalunya was founded, the first ever Catalan cultural centre to be established in America. Some of these emigrants in the 18th and 19th centuries were able to accumulate large fortunes and return to their homeland, where they built luxurious residences and came to be known as *indianos* or *americanos*. In many Latin American countries, Catalan missionaries were numerous, especially the Capuchin friars, but also other Catholic orders. At the end of the 19th century and beginning of the 20th, Costa Rica also saw the arrival of a group of Catalan traders who became very influential. Amongst these was the progressive politician Josep Figueres Ferrer, who served several terms as president of the country and who is told to have held his governmental meetings in Catalan, having several ministers also of Catalan origin. After the defeat of the Republican government at the end of the Spanish Civil War, Mexico generously received a large portion of the Catalan diaspora (Murià 1996), though significant numbers also emigrated to other countries such as Argentina (Montoya 2017), Chile, Uruguay, Colombia, and Venezuela. This big migratory wave strengthened the Catalan cultural centres in these countries, leading to numerous and remarkable publications in Catalan and even, in some cases, to the celebration of the exiled literary contest Jocs Florals de la Llengua Catalana (Manent 1992–1993).

3 The rest of the world

Africa and Asia have only sporadically received exiled or migrant Catalans in addition to religious missionaries, though contact with these continents has existed since ancient times. The Majorcan Anselm Abdullah Turmeda, a Franciscan friar who later converted to Islam, for example, was a translator and head of the customs office in Tunis at the end of the 14th and early 15th centuries. During the 19th century, after the French colonisation of Algeria, among the French and Europeans settled there, contemptuously called *pièds-noirs*, there was a large colony of people from Northern Catalonia, the Balearic Islands and Valencia (Menages/Monjo 2007). They were largely concentrated near Oran, where in the early 20th century even newspapers were being produced in Catalan, there known as *patuet* – with the French spelling. The mother of the Algerian-born French novelist, Albert Camus, for instance, was Minorcan. Because of the global wave of immigration into the Catalan territories in the late 20th century, nowadays many citizens from all continents have adopted the Catalan language in order to integrate into the host society.

Though not comparable to the presence of Catalans in Europe or in America, most recently Oceania, particularly Australia, has been the destination of a certain amount of Catalan emigration. In 1847, the Benedictine Josep M. (Benet) Serra Julià was

appointed Bishop of Victoria and, in 1926, Esteve Morell, son of a Catalan migrant and educated in Catalonia, was elected mayor of Melbourne. Today, the city of Sydney also has an active Catalan community.

4 Catalan in the contemporary world

4.1 The *Casals*: a network of Catalan communities

The *Casals*, as the associations of Catalan communities from outside the Catalan territories tend to be called – Casal de Catalunya, Casal Català or Centre Català are the most common names, but there are others, such as Casa Balear or Centre Valencià, or the Orfeó Català of Mexico – often play a very important role in the dissemination of Catalan language and culture in their respective countries. The first of these was created in 1840 in Havana (Cuba), then under the name of Societat de Beneficència dels Naturals de Catalunya. These associations have experienced ups and downs throughout their history and have served to unite the different waves of migration or exile. Currently there are far over a hundred of them, in some forty countries, spread across America, Europe, Asia and Oceania, today usually integrated into the FIEC (Federació Internacional d'Entitats Catalanes). Since 1998, the Dia Internacional de la Catalunya Exterior has been celebrated on April 24th (FIEC).

4.2 Magazines

This diaspora, whether economic or political, has produced numerous magazines, the very first having been published in Cuba in 1861. Between magazines and newsletters we know of 681 in thirty different countries. France, with 297 titles, and Mexico, with 105, are the countries with the most publications, followed by Argentina, Cuba and Chile. Amongst all this production we find mythical publications such as the weekly *L'Aureneta* (Buenos Aires, 1876–1877, 1878–1880 and 1889–1890), the monthly magazine *La Nova Catalunya* (Havana, 1908–1932), *Germanor* (Santiago de Chile, 1912–1962), the monthly publication *Ressorgiment* (Buenos Aires, 1916–1972), which was the longest-running magazine of Catalan exiles, *La Llumanera de Nova York* (New York, 1874–1881), *Quaderns de l'Exili* (Coyoacan, Mexico, 1943–1947), *La Nostra Revista* (Mexico City, 1946–1954) or *Vida Nova* (Montpellier, 1954–1978), a quarterly publication (Balcells 1988; Surroca 2004).

4.3 Radio, TV, Internet and the digital world

These Catalan communities abroad, in addition to printed publications, also produced radio programs. The first program explicitly intended “for emigrated and exiled Catalans”, *L’Hora Catalana*, began broadcasts on November 4, 1928, in Buenos Aires, and would remain on the air for 56 years. Very soon after, similar programs in Catalan were being broadcast in Chile, Uruguay, Cuba, Venezuela and Mexico. Some more ephemeral programs in Catalan were also broadcast in Santo Domingo and in Brazil (São Paulo) and, more recently, in Australia (first in Brisbane and then in Melbourne), Toronto, Paris and Toulouse. Among the international radio stations that have broadcast or are broadcasting in Catalan are Radio Moscow, the BBC, the BRT, Radio España Independiente, Radio Paris, Office of War Information (CBS, NBC), Radio Belgrade, Radio Prague, Radio Budapest, Radio Transmundial (Monaco), Radio Beijing, Radio Exterior de España, Ràdio Andorra or Ràdio de les Valls (Andorra) (Garcia-Ripoll/Niqui 2007).

Catalunya Ràdio and TV3 Televisió de Catalunya were broadcasting via satellite from 1997 to 2012 across Europe, America and Northern Africa. As of 2012 they have been broadcast by cable. However, the real revolution in the reception of television and radio in Catalan all over the world, as for so many other things, has been the Internet. Although Catalan does not yet have official recognition, it is the ninth most widely spoken language in the European Union, it has its own domain on the Internet (.cat) and, in 2017, it was the eighth language in the world in terms of online penetration, that is to say, in the number of users in relation to the total Catalan population. It is the 19th language in the world by number of web pages per speaker, the 17th by number of articles on Wikipedia and the first by number of quality articles, the 19th most used language on Twitter and the eighth most widely used language for blogs, and it is also one of the first European language to be used in machine translation.

4.4 Catalan studies abroad

4.4.1 Origins

The international expansion of Catalan studies is thought to have begun as early as the Middle Ages thanks to the growing interest sparked by the works by Ramon Llull throughout Europe. Following this came the development of bibliographical research during the 17th century. But it was not until the second half of the 19th century that we can speak of a true international dissemination of Catalan studies, with names such as Cambouliu, Helfferich, Amador de los Ríos, Menéndez y Pelayo, Cardona, Tubino, Denk and, especially, Morel-Fatio.

Catalan studies experienced a decisive boost at the very beginning of the 20th century. As a result of the enthusiasm with which Father Antoni M. Alcover inspired

his contemporaries, the first International Congress of the Catalan Language was held in Barcelona in 1906, with the participation of the great Romanists of the time, including the aforementioned Alfred Morel-Fatio, Jean-Joseph Sarrailh and Bernhard Schädcl.

Then came the works of Raymond Foulché-Delbosc, who included research on Catalan literature in his *Revue Hispanique*, the Italian contribution involving scholars such as Mario Casella, Arturo Farinelli, Ezio Levi or Bernardo Sanvisenti, the introduction of Catalan language and literature studies in Germany through the efforts of the aforementioned Bernhard Schädcl, professor at the University of Hamburg and a collaborator of Father Alcover, or the publication of certain key works such as *Das Katalanische* (1925) by Wilhelm Meyer-Lübke (Meyer-Lübke 1925). Even before the Spanish Civil War, the British tradition of Catalan studies began around Edgar Allison Peers and the *Bulletin of Hispanic Studies* at the University of Liverpool, with the collaboration of other scholars, such as William James Entwistle or Ignasi González-Llubera.

After the Second World War, a new conference, the VIIème Congrès international de linguistique romane, held in Barcelona in 1953, focused international attention on Catalan studies. In the midst of the persecution of the Catalan language and culture by Franco's regime, the Barcelona conference dedicated most of its efforts to the study of Catalan (about two-thirds of the academic contributions) and Catalan was the language used in several papers and communications as well as in discussions, in a meeting of researchers from some 20 countries.

4.4.2 Associations

Around this time, the Catalanist group in Britain (a group that had grown and incorporated Josep M. Batista i Roca, who was in exile and working as a professor at Cambridge) were preparing the creation of an entity based in London to coordinate their work and spread Catalan studies in the United Kingdom. The following year, 1954, the Anglo-Catalan Society (A-CS) was created, the oldest of the Catalan studies societies, and, from 1955, began holding an annual conference usually hosted by a British university. It used to offer an annual prize at the exiled Jocs Florals de la Llengua Catalana, and now promotes the exchange of young British and Catalan researchers through an annual scholarship. From 1980, the A-CS started *The Anglo-Catalan Society Occasional Publications* (A-CSOP); the fifteen volumes that have already appeared are freely available on the Internet (The Anglo-Catalan Society Occasional Publications). The A-CS also publishes the magazine *Journal of Catalan Studies* (JoCS), which was the first electronic journal on Catalan studies, appearing in 1988, and is also freely available on the Internet (Journal of Catalan Studies).

Little by little, and often thanks to the insistence of exiled Catalan scholars, Catalan language and literature studies have managed to make a place for themselves

in the academic programmes of European and North American universities, either by highlighting their importance in Hispanic studies or Romance studies programmes, or even, in some cases, with the creation of specific programmes.

Within this context, Professor Georges Straka, at that time director of the Centre for Philology and Romance Literature at the University of Strasbourg, offered the centre as host for a debate on Catalan entitled *La linguistique catalane*, which took place in 1968.

The success of this first conference on Catalan topics led, among other things, to a proposal for another, this time at the University of Amsterdam in 1970, where plans for a new society of Catalan studies were initiated and a managing commission was nominated to prepare a new conference and the statutes of the future society.

The third conference took place in 1973 at the University of Cambridge (UK), where the statutes were approved, the first board was nominated and the Associació Internacional de Llengua i Literatura Catalanes (AILLC) was officially created. Since then, the AILLC has held a symposium every three years and, since the end of the Spanish dictatorship, the venues have alternated between universities in the Catalan territories and institutions overseas: Basel (1976), Andorra (1979) – a conference that, due to the new political situation following the death of Franco, closed with a session at the Institut d'Estudis Catalans in Barcelona –, Rome (1982), Tarragona-Salou (1984), Toulouse (1987), Alacant-Elx (1991), Frankfurt am Main (1994), Palma, Majorca (1997), Paris (2000) – at the University of Paris-Sorbonne, which set up its Centre d'Études Catalanes back in 1977 –, Girona (2003), Budapest (2006), London (2009), Salamanca (2012), Valencia (2015), and Bucharest (2018).

The AILLC, in collaboration with several other associations, has published the half-yearly journal *Estudis de Llengua i Literatura Catalanes* since 1980.

In 1978, the Associazione Italiana di Studi Catalani (AISC) was officially created in Rome. A group of Italian scholars had been making plans for the society for a couple of years and, since the initial meeting in the Italian capital, the society has held conferences in Naples, Venice, Cagliari (Sardinia), Verona and Turin. Since 2011 it has published the yearbook *Rivista Italiana di Studi Catalani*.

In the same year 1978, but on the other side of the Atlantic Ocean, the North American Catalanists also created a new society of their own. The collaboration between exiled Catalan and American scholars had led to the creation, in 1958, of a Catalan-Provençal section – later it was split into two different sections – in the Modern Language Association Annual Conference (MLA), and to the publication of a volume in tribute to Josephine de Boer (1977), the pioneer of Catalan studies in the US. It was in 1978 that the first North American conference for Catalan studies took place at the University of Illinois at Urbana-Champaign, where the statutes were approved and the first board of the North American Catalan Society (NACS) was appointed. NACS has held conferences at the universities of Yale (New Haven, Connecticut; 1979), Toronto (Ontario, Canada; 1982), the Catholic University of Washington (Washington, DC; 1984), South Florida (Tampa, with a session in Saint Augustine, Florida;

1987), British Columbia (Vancouver, British Columbia, Canada; 1990), California (Berkeley, California; 1993), Indiana (Bloomington, Indiana; 1995), at the Institut d'Estudis Catalans (Barcelona; 1999), at Brown University (Providence, Rhode Island; 2001), Eton College (Windsor, Berkshire, United Kingdom, 2004, held jointly with the Anglo-Catalan Society to commemorate its fiftieth anniversary), Dalhousie University (Halifax, Nova Scotia, Canada; 2007), Temple University (Philadelphia, Pennsylvania; 2010); the University of Toronto (Canada; 2013), the Institut d'Estudis Catalans (Barcelona; 2015), and Indiana University (Bloomington, Indiana; 2017) again. In 1986, NACS began to publish the *Catalan Review*, the *International Journal of Catalan Culture*, which now appears yearly (Catalan Review), and since 2013 has awarded the NACS Prize for American college papers on Catalan culture.

The end of the 1970s by no means saw the end to this associative trend in international Catalan studies. 1983 saw the creation of the Deutscher Katalanistenverband (DKV), then called the Deutsch-Katalanische Gesellschaft (DKG), which organises a conference annually in a German, German-speaking Swiss or Austrian university. In 1988 it began publishing the yearbook *Zeitschrift für Katalanistik*, which is freely available on the Internet (*Zeitschrift für Katalanistik*).

In 1990, the French Catalanists also created their own organisation, the Association Française des Catalanistes (AFC), which has already held several meetings at universities in France.

In the 21st century, and after two preparatory meetings (Philadelphia, Pennsylvania, USA, 2010, and Havana, Cuba, 2011), the Associació de Catalanistes de l'Amèrica Llatina (ACAL) was established at the Colegio de Jalisco (Guadalajara, Mexico) in 2012. This is organised into national sections and has already held meetings in Mexico (Guadalajara and Puebla), Cuba (Havana), Germany (Berlin), and Catalonia (Barcelona).

All the above-mentioned academic societies make up the Federació Internacional d'Associacions de Catalanística (FIAC).

Officially forbidden in formal education in the Catalan territories, Catalan language and literature studies still had a place from 1946 onwards in the Societat Catalana d'Estudis Històrics, created as a subsidiary of the Historical-Archaeological and Philological sections of the then clandestine Institut d'Estudis Catalans. In 1986 this was transformed into a new subsidiary society, the Societat Catalana de Llengua i Literatura (SCLL), which has published the yearbook *Llengua & Literatura* (freely available on the Internet) ever since (*Llengua & Literatura*).

In addition to these Catalanist scientific publications, periodical or otherwise, published by some of these associations, there are those published by the universities and other scientific institutions in the Catalan territories, such as the Institut d'Estudis Catalans. These include journals such as *Els Marges*, *Revue d'Études Catalanes* (REC), ten issues of which were published by the Centre d'Études et de Recherches of the University of Montpellier III-Paul Valéry between 1998 and 2005, and since 2014 have been freely available on the Internet (*Revue d'Études Catalanes*); the yearbook *Estu-*

dios Catalanes, the official organ of the lectorship in Catalan of the Universidad Nacional del Litoral of Santa Fe (the Argentine Republic), which started in 2003 and of which six issues have appeared; or the monthly electronic journal *Catalonia*, published since the year 2008 by the Séminaire d'Études of the research group CRIMIC of the University of Paris-Sorbonne, freely available on the Internet (Catalonia).

4.5 The Institut Ramon Llull

The interest in Catalan language and literature shown by international academia has generated a Catalan studies network, established in difficult times by enthusiastic academics of both Catalan and non-Catalan origin from Europe and North America, despite the inexistence of a Catalan state. This network has also begun to spread to the rest of the American continent and now even exists in some places in Asia and Oceania. The resulting demand for specialised teachers for the Catalan lectorships created at the universities in these areas led the Generalitat de Catalunya (the Catalan autonomous government) to create, in 1988, a commission supporting Catalan studies abroad, a task that was assigned to the Institut Ramon Llull, a public institution created in 2002 with the aim of promoting studies of Catalan language and culture in academia abroad, the translation of literature and essays written in Catalan, and Catalan cultural production in other areas such as the theatre, cinema, circus, dance, music, visual arts, design and architecture. Currently, the IRL manages lectorships at over 150 universities around the world that teach Catalan studies, which means that, beyond the Catalan territories, more than 6,000 students every year study courses on Catalan language and culture. In addition, the IRL promotes advanced studies and research on Catalan language and culture through study centres and chairs for visiting professors at some prestigious universities, such as the Centre d'Études Catalanes at the University of Paris-Sorbonne, the Catalan Studies Research Centre at Queen Mary University of London, the Rodoreda Chair of the Graduate Center at the City University of New York, the Joan Coromines Visiting Chair of Catalan Studies at the University of Chicago, and the Juníper Serra Visiting Chair at Stanford University. In addition, several German universities offer Bachelor and Master degrees in which Catalan studies are taught, and are developing, with the support of the IRL, an additional qualification training programme in Catalan language and culture.

The IRL also organizes tests for the evaluation and certification of proficiency in Catalan as a foreign language outside of the Catalan-speaking regions. This certificate of Catalan demonstrates the level of proficiency in the Catalan language for academic and employment purposes and is divided into five levels – basic (A2), elementary (B1), intermediate (B2), competent (C1) and advanced level (C2) – in accordance with the common European framework of reference for language proficiency. At present, the tests are carried out in eighty cities across Europe, America and Oceania (IRL).

5 Bibliography and websites

- ACAL (Associació de Catalanistes de l'Amèrica Llatina), <http://acal.espais.iec.cat/> (last accessed: 30.03.2017).
- A-CS (Anglo-Catalan Society), <http://www.anglo-catalan.org/> (last accessed: 30.03.2017).
- A-CSOP (The Anglo-Catalan Society's Occasional Publications), <http://www.anglo-catalan.org/publications.html> (last accessed: 30.03.2017).
- AFC (Association Française des Catalanistes), <http://france-catalaniste.com/> (last accessed: 30.03.2017).
- AILLC (Associació Internacional de Llengua i Literatura Catalanes), <http://aillc.espais.iec.cat/> (last accessed: 04.11.2018).
- AISC (Associazione Italiana di Studi Catalani), <http://aisc.llocs.iec.cat/> (last accessed: 30.03.2017).
- Balcells, Josep Maria (1988), *Revistes dels catalans a les Amèriques*, Barcelona, Generalitat de Catalunya.
- Bover i Font, August (1993), *Manual de catalanística*, Barcelona, Publicacions de l'Abadia de Montserrat/Diputació de Tarragona.
- Bover i Font, August (2007), *Sardocatalana. Llengua, literatura i cultura catalanes a Sardenya*, Paiporta, Denes.
- Carbonell, Jordi (1984), *La llengua i la literatura medieval i moderna*, in: Jordi Carbonell/ Francesco Manconi (edd.) (1984), *Els catalans a Sardenya*, Barcelona, Enciclopèdia Catalana/Generalitat de Catalunya/Consell Regional de Sardenya, 93–98/*La lingua e la letteratura medievale e moderna*, in: Jordi Carbonell/Francesco Manconi (edd.) (1984), *I catalani in Sardegna*, Cinisello Balsamo, Consiglio Regionale della Sardegna/Generalitat de Catalunya, 93–98.
- Casanova, Eugeni (2014), *La tribu perduda dels catalans: els gitanos catalans de França*, in: Joan A. Argenter (ed.), *Enclavaments lingüístics i comunitats locals. El català a Sardenya, el Carxe i entre els gitanos catalans de França*, Barcelona, Institut d'Estudis Catalans, 221–242.
- Casanova, Eugeni (2016), *Els gitanos catalans de França. Llengua, cultura i itineraris de la gran diàspora*, Lleida, Pagès.
- Catalan Review, <http://online.liverpooluniversitypress.co.uk/loi/catr> (last accessed: 30.03.2017).
- Catalonia, <http://crimic-sorbonne.fr/publication-crimic/8681/> (last accessed: 30.03.2017).
- Coromines, Joan (1954), *El que s'ha de saber de la llengua catalana*, Palma, Moll.
- Crespo i Sala, Martí (2018), *Els "minorkeens" de Gibraltar*, Barcelona, Publicacions de l'Abadia de Montserrat/Institut Menorquí d'Estudis.
- DKV (Deutscher Katalanistenverband), <http://www.katalanistik.de/> (last accessed: 30.03.2017).
- Escudero, Jean-Paul (2014), *Enclavaments lingüístics dels gitanos catalanoparlants al Rosselló i a Occitània: un estat de la qüestió*, in: Joan A. Argenter (ed.), *Enclavaments lingüístics i comunitats locals. El català a Sardenya, el Carxe i entre els gitanos catalans de França*, Barcelona, Institut d'Estudis Catalans, 209–219.
- FIAC (Federació Internacional d'Associacions de Catalanística), <http://fiac.espais.iec.cat/> (last accessed: 30.03.2017).
- FIEC (Federació Internacional d'Entitats Catalanes), <http://www.fiecweb.cat/> (last accessed: 30.03.2017).
- Garcia-Ripoll Duran, Martí/Niqui Espinosa, Cinto (2007), *La ràdio en català a l'estranger*, Bellaterra, Publicacions de la Universitat Autònoma de Barcelona.
- IRL (Institut Ramon Llull), <http://www.llull.cat/catala/home/index.cfm> (last accessed: 30.03.2017).
- Journal of Catalan Studies, <http://jocs.anglo-catalan.org/ojsnew/index.php/jocs/index> (last accessed: 30.03.2017).
- Llengua & Literatura, <http://revistes.iec.cat/index.php/LLiL/index> (last accessed: 30.03.2017).

- Manent, Albert (ed.) (1992–1993), *Diccionari dels catalans d'Amèrica*, 4 vol., Barcelona, Comissió Amèrica i Catalunya 1992/Generalitat de Catalunya.
- Menages, Àngela-Rosa/Monjo, Joan-Lluís (2007), *Els valencians d'Algèria (1830–1962): memòria i patrimoni d'una comunitat emigrada*, Picanya, Edicions del Bullent.
- Meyer-Lübke, Wilhelm (1925), *Das Katalanische*, Heidelberg, Winter.
- Montoya Abad, Brauli (2017), *La llengua dels mallorquins de San Pedro (Argentina)*, Barcelona, Publicacions de l'Abadia de Montserrat.
- Murià, José María (ed.) (1996), *Diccionario de los catalanes de México*, Zapopan, Colegio de Jalisco/Generalitat de Catalunya.
- NACS (North American Catalan Society), <http://nacs.espais.iec.cat/> (last accessed: 30.03.2017).
- Rasico, Philip D. (1990), *The Minorcans of Florida: Their History, Language and Culture*, New Smyrna Beach (Florida), Luthers.
- Revue d'Études Catalanes, <http://france-catalaniste.com/documentation/rec#NOUVELLE> (last accessed: 30.03.2017).
- SCLL (Societat Catalana de Llengua i Literatura), <http://blogs.iec.cat/scll/> (last accessed: 30.03.2017).
- Surroca, Robert (2004), *Prensa catalana de l'exili i l'emigració (1861–1976)*, Barcelona, Generalitat de Catalunya.
- Vallverdú, Francesc (ed.) (2001), *Enciclopèdia de la llengua catalana*, Barcelona, Edicions 62, 326–369.
- Zeitschrift für Katalanistik, <http://www.romanistik.uni-freiburg.de/pusch/zfk/dt/index.htm> (last accessed: 30.03.2017).

List of Contributors

Joan A. Argenter

Emeritus Professor of General Linguistics
Universitat Autònoma de Barcelona
Facultat de Filosofia i Lletres
Departament de Filologia Catalana
Edifici B – Campus de Bellaterra
E-08193 Cerdanyola del Vallès
joan.argenter@uab.cat

Fellow of the
Institut d'Estudis Catalans
Philology Section
Carrer del Carme 47
E-08001 Barcelona
jargenter@iec.cat

UNESCO Chair on Cultural and
Linguistic Diversity
Chairholder
Institut d'Estudis Catalans

Montserrat Bacardí

Associate Professor of Translation and
Contemporary Catalan Literature
Universitat Autònoma de Barcelona
Facultat de Traducció i d'Interpretació
Edifici K
E-08193 Bellaterra, Cerdanyola del Vallès
montserrat.bacardi@uab.cat

Emili Boix-Fuster

Professor of Catalan Philology
Universitat de Barcelona
CUSC – Centre Universitari de Sociolingüística i
Comunicació
Gran Via de les Corts Catalanes 585
E-08007 Barcelona
eboix@ub.edu

August Bover i Font

Emeritus Professor of Catalan Philology
Universitat de Barcelona
Departament de Filologia Catalana i

Lingüística General
Gran Via de les Corts Catalanes, 585
E-08007 Barcelona
bover@ub.edu

Fellow of the Institut d'Estudis Catalans

Jenny Brumme

Professor of Translation and Applied Linguistics
Universitat Pompeu Fabra
Department of Translation and Language
Sciences
c/ Roc Boronat 138
E-08018 Barcelona
jenny.brumme@upf.edu

M. Teresa Cabré

Professor Emerita of Catalan Philology
Universitat Pompeu Fabra
Institute for Applied Linguistics
Ciutadella Campus
Roc Boronat, 25–27
E-08005 Barcelona
teresa.cabre@upf.edu

Fellow of the Institut d'Estudis Catalans
Chairman of the Philology Section

Oriol Camps Giralt

Fellow of the
Institut d'Estudis Catalans
Philology Section
Carrer del Carme 47
E-08001 Barcelona
oriol.camps@allerats.cat

Former language advisor of Catalunya Ràdio
(1987–2013)
Av. Diagonal 614
E-08021 Barcelona

Former language advisor of Corporació Catalana
de Mitjans Audiovisuals (2007–2013)
c/ Jacint Verdaguer, 81
E-08970 Sant Joan Despí, Barcelona

<https://doi.org/10.1515/9783110450408-034>

Maria Josep Cuenca

Professor of Catalan Philology
 Universitat de València
 Institut Interuniversitari de Filologia Valenciana /
 Dept. de Filologia Catalana
 Facultat de Filologia, Traducció i Comunicació
 Avda. Blasco Ibáñez 32, 7a
 E-46010 València
 maria.j.cuenca@uv.es

Fellow of the Institut d'Estudis Catalans

Nicolau Dols

Associate Professor of Catalan Philology
 Universitat de les Illes Balears
 Departament de Filologia Catalana i Lingüística
 General
 Edifici Ramon Llull
 Carretera de Valldemossa, km. 7,5
 E-07122 Palma (Illes Balears).
 n.dols@uib.cat

Fellow of the Institut d'Estudis Catalans

PD Dr. Ingo Feldhausen

Goethe-Universität Frankfurt
 Institut für Romanische Sprachen und Literaturen
 Norbert-Wollheim-Platz 1
 D-60629 Frankfurt am Main
 Deutschland
 feldhausen@em.uni-frankfurt.de

Antoni Ferrando Francés

Professor of Catalan Philology
 Universitat de València
 Departament de Filologia Catalana
 Facultat de Filologia,
 Avda. Blasco Ibáñez 32
 E-46010 València
 antoni.ferrando@uv.es

Fellow of the Institut d'Estudis Catalans

Joan Julià-Muné

Professor of Applied Phonetics
 Universitat de Lleida
 Department of Catalan Philology and
 Communication
 Pl. Víctor Ciurana 1

E-25003 Lleida
 jjulia@filcat.udl.cat

Dr. Aina Labèrnia Romagosa

Senior lecturer
 Universitat Ramon Llull
 Translation and Language Sciences Department
 Santaló, 37
 E-08021 Barcelona
 alabernia@peretarres.url.edu

Xavier Lamuela

Professor of Romance Linguistics, retired
 Universitat de Girona
 Departament de Filologia i Comunicació
 Campus Barri Vell
 Pl. Ferrater Mora, 1
 E-17004 Girona

Jens Lüdtke (†)

Professor of Romance Philology
 Universität Heidelberg
 Romanisches Seminar
 Seminarstraße 3,
 D-69117 Heidelberg

Joaquim Mallafre

Professor of Catalan Philology, retired
 Universitat Rovira i Virgili
 Facultat de Lletres – Campus Catalunya
 Avinguda Catalunya 35
 E-43002 Tarragona

Fellow of the Institut d'Estudis Catalans
 jmallafre@iec.cat

Josep Martines

Professor of Catalan Philology
 Universitat d'Alacant
 Departament de Filologia Catalana
 Carretera de Sant Vicent del Raspeig
 E-03690 Sant Vicent del Raspeig
 josep.martines@ua.es

Fellow of the Institut d'Estudis Catalans

Mar Massanell i Messalles

Associate Professor of Catalan Philology
 Universitat Autònoma de Barcelona
 Departament de Filologia Catalana
 Facultat de Filosofia i Lletres
 Campus de Bellaterra
 E-08193 Bellaterra (Cerdanyola del Vallès)
 mar.massanell@uab.cat

Jaume Mateu

Associate Professor of Catalan Philology
 Universitat Autònoma de Barcelona
 Departament de Filologia Catalana
 Facultat de Filosofia i Lletres
 Edifici B
 Carrer de la Fortuna, s/n
 E-08193-Bellaterra (Cerdanyola del Vallès)
 jaume.mateu@uab.cat

Joan Francesc Mira

Lecturer of Greek Philology, retired
 Universitat Jaume I de Castelló
 Facultat de Ciències Humanes i Socials
 Departament de Filologia i Cultures Europees
 Av. de Vicent Sos Baynat s/n
 E-12071 Castelló de la Plana

Fellow of the Institut d'Estudis Catalans
 Philosophy and Social Sciences

Founder and former director of the Valencian
 Museum of Ethnology

M. Amor Montané

Researcher on Terminology and Applied
 Linguistics
 Universitat Pompeu Fabra
 Institut for Applied Linguistics
 Poblenou Campus
 Ramon Trias Fargas 25–27
 E-08005 Barcelona
 amor.montane@upf.edu

Miquel Nicolás

Associate Professor of Catalan Philology
 Universitat de València
 Departament de Filologia Catalana
 Av. Blasco Ibáñez, 32
 E-46010 València
 nicolas@uv.es

Manuel Pérez Saldanya

Professor of Catalan Philology
 Universitat de València
 Departament de Filologia Catalana
 Facultat de Filologia, Traducció i Comunicació
 Avinguda Blasco Ibáñez, 32
 E-46010 València
 saldanya@uv.es

Fellow of the Institut d'Estudis Catalans

Eva Pons Parera

Associate Professor
 Universitat de Barcelona
 Departament de Ciència Política, Dret
 Constitucional i Filosofia del Dret
 Av. Diagonal, 684
 E-08028 Barcelona

Editor of *Revista de Llengua i Dret*/
Journal of Language and Law

Miquel Àngel Pradilla Cardona

Associate Professor of Catalan Philology
 Universitat Rovira i Virgili
 Facultat de Lletres – Campus Catalunya
 Avinguda Catalunya, 35
 E-43002 Tarragona

Fellow of the Institut d'Estudis Catalans

Director of the Cruscat Network
 Institut d'Estudis Catalans

President of the Catalan Society
 of Sociolinguistics

Joan Pujolar

Professor of Sociolinguistics
Universitat Oberta de Catalunya
Estudis d'Arts i Humanitats
Av. Tibidabo 39–43
E-08035 Barcelona

August Rafanell

Associate Professor of Catalan Philology
Universitat de Girona
Institut de Llengua i Cultura Catalanes
Pl. Ferrater Mora, 1
Campus Barri Vell
E-17004 Girona
august.rafanell@udg.edu

Philip D. Rasico

Professor of Spanish and Catalan Language
and Linguistics
Vanderbilt University
Department of Spanish and Portuguese
Nashville, Tennessee 37240-1617
U.S.A
philip.d.rasico@vanderbilt.edu

Gemma Rigau

Professor of Catalan Philology
Universitat Autònoma de Barcelona
Departament de Filologia Catalana
Edifici B. Campus UAB
E-08193 Cerdanyola del Vallès
gemma.rigau@uab.cat

Fellow of the Institut d'Estudis Catalans

Joan Soler Bou

Institut d'Estudis Catalans
Carrer del Carme, 47
E-08001 Barcelona
jsoler@iec.cat

Albert Turull

Associate Professor of Catalan Philology
Universitat de Lleida
Departament de Filologia Catalana i
Comunicació
Plaça de Víctor Siurana 1
E-25002 Lleida
turull@filcat.udl.cat

Fellow of the Institut d'Estudis Catalans

F. Xavier Vila

Associate Professor of Catalan Philology
Departament de Filologia Catalana i Lingüística
General
CUSC – Centre de Recerca en Sociolingüística i
Comunicació
Universitat de Barcelona
Gran Via de les Corts Catalanes, 585
E-08007 Barcelona
fxvila@ub.edu

Xavier Villalba

Associate Professor of Catalan Philology
Universitat Autònoma de Barcelona
Departament de Filologia Catalana
Carrer de la Fortuna s/n
Campus UAB
E-08193 Cerdanyola del Vallès
Xavier.Villalba@uab.cat

Kathryn A. Woolard

Research Professor/Professor Emerita
of Anthropology
UC San Diego
Department of Anthropology, 0532
9500 Gilman Drive
La Jolla, CA 92093-0532
kwoolard@ucsd.edu

Edward Sapir Book Prize winner 2017

Index

- absolute clauses (also: absolute participle construction) 137–138, 194, 227–228, 234, 236
- Academia* 492, 499, 538
- Acadèmia* 50, 61, 62, 413, 414, 428, 492, 507, 518, 520, 533, 534, 548, 559, 686, 715
- academy 61, 316, 414, 416, 507, 519, 527, 562, 657, 676
- Academy 32, 40, 59, 62, 63, 414, 507, 660, 661, 686, 699, 714
- accent 607
 - *Xava* accent 407
- accent (see also: stress) 85, 359
- accent marks 85–92, 107, 616
- Acció Cultural del País Valencià 78, 678
- accusative 137–138, 156, 176–177, 217–218, 386, 458–461
- ací* (locative adverb) 154, 189
- açò* (demonstrative pronoun) 154, 171, 180
- acoustic 404–407, 452
- activist(s) 39, 533, 602, 709, 710, 724, 727, 728, 729
- addressee 133–153, 167, 171, 189, 248, 288, 291, 298, 301–306, 323–324, 608–611
- addresser (also: addressor) 133, 134, 139, 153, 171, 189, 288, 291, 295, 298, 299, 301, 307
- adjectival function 133
- adjectivalisation 358–359
- adjective 122, 131, 132, 135–138, 142, 144, 147–150, 176–177, 180–181, 182, 186–191, 215–216, 230–233, 239, 264, 299, 329–333, 358–359, 361, 364, 366, 367, 380, 389, 498, 568, 569, 575, 576, 625
- adjective phrase (AdjP) 168, 282
- adjunct 131, 166, 168–169, 179, 184, 194–195, 220, 226–227, 235, 236–237, 239–240, 241, 255, 259, 279–280, 356–358, 367
- administration 312, 315–316, 335, 337, 382, 392, 411, 478, 479, 482, 528–529, 536, 540, 546, 552, 630, 635, 642, 643, 653–666, 672, 730
- administrative division 374, 388, 439, 637, 652
- adstrate (also: adstratum) 388, 567, 573
- adverb 115, 119, 122, 131, 132, 166, 177, 180, 182, 186–191, 193, 206, 207, 208, 212, 213, 214, 215, 218, 219, 220–223, 227, 230, 231–233, 240, 241–242, 259, 262, 266, 291, 299–331, 357, 359, 360, 361, 362, 367, 385, 462, 474
- adverb phrase (AdvP) 168, 282
- adverbial locution 187–188
- adverbialisation 359
- advocate(s) 501, 511–512, 716
- Aebischer, Paul 561, 565, 577
- affirmative polarity 206–208
- affix 145
 - derivational affix 130, 329, 331, 352, 354, 365
 - inflectional affix 130, 142, 146, 565
- age group(s) (variable) 342, 404–409, 427, 427, 609, 640
- agent 184, 194, 205, 279, 353
- agglutinated, agglutination 107, 108, 388, 465, 568, 572, 576
- agreement
 - gender agreement 137–138, 170
 - number agreement 137–138, 141, 170, 386
 - past participle agreement 42, 135–136, 137–138, 465–466
 - person agreement 138, 141
- Aguiló, Marian 39, 42, 46, 47, 504, 524, 528
- Alacant 12, 189, 408, 524, 531, 563, 610, 622, 632, 633, 634, 638
- Alacantine 12, 383–384, 623
- Alcover, Antoni M. 39–47, 50–56, 61, 76–77, 377, 379, 380–383, 414, 506, 524–527, 532, 533, 535–538, 549, 555, 562, 744–745
- l'Alguer (It. Alghero) 9, 12, 315, 318, 328, 335, 340, 373–377, 401, 412, 449, 451, 473–475, 527, 532, 570, 599, 608, 616–621, 630–646
- Alguerese 12, 87, 104, 110–124, 146, 152, 155–163, 373, 377–390, 619–622, 645, 656, 675
- allomorph, allomorphic, allomorphism 145–146, 159–162, 362, 624
- Almirall, Valentí 3, 505, 507–509, 520–521
- alphabetic (see: orthography, spelling, writing)
- analogical form 144, 391, 460, 462, 623
- analogy 82–83, 96, 116, 148–150, 151, 155–157, 389, 456
- anaphoric 132, 154, 175, 177, 190, 221, 240–241, 256, 266
- Ancien Régime 617, 631–633

<https://doi.org/10.1515/9783110450408-035>

- Andalusi* Arabic 478, 599
Andalusi Romance 314–315, 446, 567
 Andorra 9, 38, 182, 371–374, 401, 412, 438–439, 473, 553, 616, 618, 634–646, 649, 651, 653–655, 674–675, 679, 686, 713, 744, 746
 Anglo-Catalan Society (A-CS) 745–747
 animate 178, 183, 196, 197, 220, 280, 356,
 antecedent 175, 215–216, 219–222, 224, 232, 256, 293, 298, 437
 anthropological, anthropology 30–31, 402, 410, 724, 726
 apocope 380, 389, 442, 451–452
 apostrophe 85, 87, 93, 96–97
 apparent time 404, 408
 appropriateness 288, 416, 530, 684, 686, 688–690, 701
 appropriation of another's voice 613–614
aquest (demonstrative) 119, 153, 171, 291–296, 385–386, 391, 459
 Aquitanian 313
 Arabic 19, 88, 110, 314, 325, 335, 443–444, 476, 478–479, 481, 566–567, 569, 573, 575–577, 599 630–631, 733
 Arabs 443–444, 565–566, 576
 Aracil, Lluís V. 60, 602, 715
 Aragon 9, 19, 87, 292, 315, 332–333, 414, 445, 471, 486–487, 499, 510, 527, 536, 579, 599–600, 616, 618, 630, 642, 651, 659–660, 674, 739–741
 – Council of Aragon 479
 – Crown of Aragon (also: Kingdom of Aragon) 315, 318, 445, 471–479
 Aragon strip (see: la Franja)
 Aragonese 42, 90–91, 315, 333–335, 372–374, 414, 439, 472–479, 599, 659–660
 Aramon i Serra, Ramon 53, 555–556
 Aran, Valley 439
 – Occitan-speaking domain 439, 473
 archaic 329, 334–335, 380, 389, 475, 520–521, 528, 530, 532, 547, 559, 567–568
 Argentina 51, 611, 742–743
 argument of the verb 192–195, 215–216, 247, 255, 259, 280, 351
 – external argument 192–194, 351
 – internal argument 192–195, 247, 280
 Arteaga Pereira, Josep M. 45
 article
 – definite article 84, 93, 107, 117, 132, 135, 138, 152–153, 170–173, 182–183, 186, 211, 220, 288, 294, 296–297, 308, 385, 387–388, 390–391, 458–460, 474, 568, 599, 623
 – definite articles of Old Catalan 458
 – indefinite article 117, 132, 137–138, 147, 153, 170–173
 – individualizer or neuter article 171, 234
 – personal article 93, 153, 171, 288–289, 387–388, 391
 aspect 136, 139–141, 146, 158, 181, 188, 190, 196, 198–199, 201, 204, 228, 277, 361, 363, 624
 assimilation (see: consonants)
 assimilation, assimilate (social assimilation) (see also: assimilation policies) 538, 632
 Associació dels Catalanistes de l'Amèrica Llatina (ACAL) 747
 Associació Internacional de Llengua i Literatura Catalanes (AILLC) 52, 78, 746
 Association Française des Catalanistes (AFC) 747
 Associazione Italiana di Studi Catalani (AISC) 746
 Asturian-Leonese 442
Atles Lingüístic del Domini Català 56, 62, 378
 augmentatives 103, 115, 360–361
 Autonomous Community 8, 371–372, 415, 649–651, 661, 666, 717
 – of Aragon 371–372, 412, 414, 644, 659
 – of Catalonia 371–372, 616
 – of Murcia 372, 658
 – of the Balearic Islands 372, 414, 663
 – of Valencia 372, 414, 660–661
 Badia i Margarit, Antoni M. 48, 50–55, 56–60, 78, 552–553, 554–557, 715
 Bakhtin, Mikhael 612
 Balari, Josep 39, 47, 512, 521
 Balearic 7, 87, 92–95, 102, 104–106, 113–114, 117, 120, 124, 146, 151–153, 157, 160–163, 253, 259, 373–374, 376–380, 383, 386–390, 474–475, 492, 568, 686
 Balearic Islands 9, 12, 39, 52, 53, 289, 300, 323, 328, 331–333, 335–336, 339, 372–374, 389, 412–414, 451, 454, 458, 472, 474, 475, 478, 489, 511, 518, 524, 531, 547, 553–554, 564, 566, 581, 616, 618, 630, 634–637, 638–646, 651, 655, 662–663, 673–674, 677, 679, 686, 699, 726, 742
 Ballot i Torres, Josep Pau 38, 491, 501, 512, 517–519

- Barcelona
- House of (also: Count of) 19, 315, 471, 472–473, 600
- Barnils, Pere 45–48, 50, 529
- Bases de Manresa* (1892) 505, 508–509
- Basque 5, 10, 21, 24, 41, 88–90, 151, 313, 438–439, 442, 563, 578
- Bastardas, Joan 441, 446, 669
- Bastero, Antoni de 38
- Berbers 443
- Bible 6, 43, 289, 478, 481, 583, 587
- Bible translations 6–8, 478, 583, 587
- bilingual
- conversation 711
 - education (see: education)
 - speakers 598, 601, 603, 605, 615
- bilingualisation (also: bilingualization) 316, 344, 617, 633, 662
- bilingualism 423, 479, 533, 598, 601, 603, 606, 645, 666, 672–673, 711–712, 715–716
- bivalency 615–616
- Bofarull, Antoni de 38
- border 8–9, 19, 30, 95, 112, 336, 382, 590, 632, 634, 658, 694, 700–701, 725
- Borja, Alfons 473
- Borja, Roderic 473, 483
- borrowing 598, 608, 612
- boundary 21, 25, 28, 84, 123, 255–256, 332, 360, 365, 372, 374, 382, 402, 571, 586, 612, 712–715, 725
- Bourbon, Bourbons 31, 487, 499, 600, 631
- Bressola, la* 618, 657, 675
- Britain 27, 45, 51, 332, 745
- Butlletí de Dialectologia Catalana* 47, 51, 54, 529, 699
- Calveras, Josep 45, 534
- Capmany, Antoni de 492, 500
- Carolingian 314, 442, 445–446, 565, 577
- Carthage 313, 437
- Carxe, el 372, 374, 616, 658
- Casas-Carbó, Joaquim 39, 522, 523, 525, 587
- Castile
- Council of Castile 487
 - Kingdom of Castile (also: Crown of Castile) 31, 316, 472–473, 478, 600, 630
- Castilian (see also: Spanish)
- literature 29, 32–33, 316
 - Monarchy (also: Spanish) 31, 33, 535, 600
- Castilian (language) (see also: Spanish)
- an Ibero-Romance language 314, 599
 - anonymity 711
 - Castilian-speaking 634, 660, 672, 713
 - Court language 473–478, 482, 583
 - education in Castilian 664
 - hegemony 634, 660, 672, 713
 - imposition 31, 316, 487, 670
 - influence 189, 198
 - interference 227, 317, 319, 331, 336, 411, 413, 511, 622, 624, 701
 - language of administration 670
 - language of the Court 666
 - language of the Crown 316
 - lexicon 315–317
 - means of instruction 671–672
 - monolingualism 479, 642
 - national language 488, 511, 633, 635, 645
 - official language 554, 631, 633, 644, 651, 660
 - official status 554, 660, 661, 662, 663–664, 677
 - orthography 499, 520
 - prestige 315–316, 476
 - reading 33, 638, 684
 - schooling 33, 670
 - speakers 642, 663
 - state language 646, 649, 651
 - teaching 492, 669, 674, 678
 - usage 317
 - varieties 315, 317
 - vehicular language 672
 - writing 33
- castilianization 537, 631–632, 676, 713
- català*
- *català acadèmic* 521
 - *català light* vs *català heavy* 604, 714
 - *català que ara es parla* 512
 - origin of the term 438, 476
- Catalan
- Catalan-Aragonese confederation 471, 600
 - Catalan *Casals* 741, 743
 - Catalan Countries 12, 20, 24, 617, 631, 635–636, 639, 643, 687, 739
 - Catalan-speaking community 8–11, 28, 518, 634–635, 640, 666, 698, 700, 743
 - Catalan Studies abroad 46, 55, 741, 744–748

- culture 28, 31, 42, 54–55, 318, 482, 504, 519, 552, 588, 632, 655, 710, 714, 726, 747
- institutions 528, 670, 729
- linguistics 1, 3, 5, 10–11, 37–38, 40, 45–47, 51–52, 54, 56, 59–61, 78, 287, 365–366, 380, 411, 422, 425, 529, 701
- magazine(s) 54, 498, 506, 521, 523–525, 531–532, 534–539, 546, 550–553, 558, 585, 614, 743, 745
- medium-sized language community 645
- publications 540, 553, 589
- Catalan (language)
 - a ‘medium-sized European language’ 2, 636, 645, 649
 - acquisition as an L2 399, 711
 - and Occitan 7, 88–92, 318, 322, 335, 338, 442, 459–460, 474–476, 480, 540–541, 567, 573, 599
 - anthroponymy (see also: personal names) 576–579
 - as a foreign language 59, 674, 680, 703
 - at the papal court in Rome 473
 - Catalan *scripta* 471, 474, 490–491
 - Catalan-speaking area 1, 12, 37, 38, 42, 46, 48, 51, 52, 55
 - Catalan-speaking enclave 9, 373, 599
 - Catalan-speaking strip (see: *la Franja*)
 - Catalan-speaking territories 51, 55, 60, 288, 476, 492, 517–519, 524, 532, 545–546, 551, 553, 558, 576, 582, 630–636, 639, 641–642, 646, 665, 670, 676, 709, 716–717, 732
 - Central Catalan 12, 113–115, 118, 120, 146, 155, 160, 162, 164, 251–253, 300, 373, 377–378, 380–383, 387–390, 623, 687
 - Chancery language (also: Chancery koine) 334, 477–482, 486, 581–582
 - codification 5, 9, 41–45, 49, 82–96, 415, 482, 504, 512, 517, 687
 - colloquial Catalan 59, 111, 113, 139, 144, 222, 328, 337–338, 340, 401, 412, 426, 475, 481
 - common Catalan 555, 608, 611, 715
 - constitutive Romance language 440
 - continental Catalan 380, 475
 - cultured language 28–29, 477, 549
 - descriptive grammar 39–40, 44, 57
 - descriptive linguistics 38, 41, 44
 - descriptive phonetics 48, 53
 - dialects 9, 48, 56, 63, 82, 87, 90, 91, 144–146, 162, 189, 292, 300, 329, 334, 371 375, 380–382, 384 386–387, 389, 450, 457, 492, 511, 526 (see also: Alguerese, Balearic, Eastern Cat., Eivissenc, Eivissenc-Formenteran, Majorcan, Minorcan, Northern Cat., Northeastern Cat., Northwestern Cat., Valencian, Western Cat.)
 - Early Catalan (see also: Old Catalan) 442, 454, 466
 - Eastern Catalan 13, 102, 144, 146, 148–151, 155, 160, 162, 327–328, 371, 373–379, 383, 386–387, 389–390, 399, 406, 415, 450, 454–455, 474–475, 491, 511, 522, 615–616
 - Eastern-Pyrenean Catalan 380
 - elocution and orthology 54, 58, 63, 540
 - fragmentation 398, 411–412, 415, 476, 501, 511
 - grammar 39, 46, 49, 51, 131, 146, 152, 414, 482, 525, 537
 - Gypsy Catalan, *Gitano* 614, 621
 - historical grammar (see also: Old Catalan) 51, 55–56, 148, 312, 522
 - language of identification 412, 639, 643–644
 - language teaching 678
 - language unity 8–9, 343, 415, 477, 497, 493, 524, 527, 558–562, 715
 - learning 675
 - lexical hispanization of Catalan 475
 - lexical (semantic) variation 332–333, 339, 341
 - *lingua franca* 413, 476, 732
 - linguistic community (also: language community) 29–30, 374, 397–400, 412, 414–415, 417, 476, 510, 619, 632, 638–640, 645, 691, 712–713
 - linguistic domain (see also: Catalan-speaking area) 147, 315, 335–336, 339, 378, 387–388, 562, 567
 - literary Catalan 39, 519–522, 530, 548
 - literary use 480–482, 503, 511
 - Mediaeval/Medieval Catalan 322, 437–467, 474–477
 - Modern Catalan 5–6, 42, 150, 157, 320, 379, 458, 485–493
 - name of the language 19, 531, 558–559, 715
 - national language 25, 536, 546, 584

- neology 704
- Northeastern Catalan 39, 332, 474, 477
- Northern Catalan (also: Rossellonese) 13, 54, 102, 105, 113–114, 117, 124, 146, 151, 155, 157, 160, 162, 163, 300, 317, 343, 373–375, 380–383, 388–390
- Northwestern Catalan 8, 13, 110–111, 113–114, 116, 146, 152, 157, 160, 162–164, 184, 189, 239, 318, 332, 373, 375, 378, 380–387, 390, 407, 454, 456–458, 474, 688
- obsolescent Catalan 611, 622–624
- official language 492, 507–508, 540–541, 600, 618, 634, 649, 650, 651, 653–654, 665, 677
- official status 490, 540, 545, 554, 646, 649–651, 653, 660–661, 662, 663, 665
- onomastics 561–579
- orthography 11, 48–49, 76, 81–82, 87–97, 107, 148, 520, 521, 536, 607
- phonetics 43, 47, 48, 56–57
- Preliterary Catalan 147, 374, 380
- purism 714
- Pyrenean Catalan 42, 380
- schooling 604, 606, 638
- semantic erosion 317
- sociolinguistics 2, 9, 10, 60, 287, 397–399, 400, 403, 633, 716, 726
- speakers 30, 316, 344, 382, 401, 407–408, 426, 475, 545, 612, 619, 620, 631, 633, 636, 638–639, 643–646, 659, 678, 711–712, 713, 727–728, 732, 741
- spelling 6–8, 11, 38, 43, 52, 83, 87–90, 93–97, 107, 114, 118, 148–149, 358, 386, 409–410, 456, 491, 501–502, 506, 511–512, 519–521, 522–523, 525, 527, 528–529, 531, 533–534, 537, 548, 550, 568, 613, 733
- standard language 6, 458, 517, 522, 555, 686
- standard variety 329, 398, 399, 410, 416, 417, 458, 605, 687–690
- target language 294, 297, 582, 588, 624
- teaching 59, 638, 678–679
- terminology 2–3, 695, 698–705
- toponym, toponymy (see also: place names) 12, 314, 442–446, 448–452, 457–458, 562–576
- transitional Northern Catalan 382–383
- trans-Pyrenean 9, 487
- uniform language 29, 490
- variation and change 312, 319, 320
- variety, varieties 25, 39, 54, 144–146, 153–155, 252, 288–289, 300, 317, 321, 332–343, 371–388, 531, 614, 619, 656, 675, 690, 704
- vocabulary 4, 39, 53, 311–343, 475, 481, 503, 530, 622, 685, 740
- Western Catalan 13, 104, 144, 146, 148–151, 155, 157, 160, 162, 315, 332, 335–336, 373–380, 384, 386–387, 390, 406, 474, 511
- Western-Eastern division 374, 376, 378
- writing 331, 412–413, 415, 456, 490, 491, 492, 501, 503, 518, 520, 521, 531, 537, 555, 557, 591, 636, 685, 689, 698, 703
- written language 28, 31, 446, 475, 477, 522, 527, 532, 548, 600, 685, 688
- Catalan literature 17, 33, 476, 481, 497, 519, 581, 710
- Catalan worldwide 616, 739–748
- Catalanism 3, 506–510, 532, 535–536, 538, 540–541, 546, 586, 710, 740
- catalanization (also: catalanisation) 530, 558, 567, 711, 713, 714, 740
- Catalanness 25, 603, 614–615, 690
- Catalonia
 - French Catalonia 6–7, 519, 524, 535, 548, 572
 - Northern Catalonia (also: *Catalunya del Nord*) 6, 13, 315, 318, 328, 333, 335, 339, 372–375, 377, 382, 401, 412, 565, 583, 614, 616–618, 631, 634–635, 637–641, 643, 645–646, 650, 653, 657, 670, 675–676, 679, 711, 715, 742
 - ‘New Catalonia’ (also: *Catalunya Nova*) 474, 476, 566, 573
 - ‘Old Catalonia’ (also: *Catalunya Vella*) 442–444, 455, 474, 566–567
 - Principality of (also: *Principat de Catalunya*) 28, 31, 455, 471–472, 486, 490, 600
 - self-government 12, 506, 603–604, 664
- Catalunya
 - *Generalitat de Catalunya* 2, 40, 48–49, 77–78, 540, 612, 656, 664, 670, 727, 729, 748
 - *Mancomunitat de Catalunya* 3, 40–41, 49, 77, 81, 528–538, 601, 670
 - origin of the term 438
- Catholic 25, 471, 473–474, 477, 482, 486, 489, 509, 553, 577, 674, 710, 729, 731, 742, 746
- Celtiberians 438
- Celtic 313–314, 438, 564, 573

- lexical-semantic change 320, 334–335, 339, 344
 - morphological change 381
 - phonetic change 62, 390
 - semantic change 7, 320–325, 339
- channel 400
- Church (see also: Catholic) 28–29, 31, 316, 339, 344, 446, 481, 489, 524, 551–553, 631, 710, 713, 724
- Church Slavonic 31
- classification 83, 102, 107, 161, 188–189, 271, 273, 283, 337, 352, 354, 360, 383, 406, 425, 574, 579, 588, 637, 690
- clause(s) 180, 207, 227, 235, 236, 238, 240, 242, 243, 251, 254, 260, 280
 - absolute (see: absolute clauses)
 - causal 226, 236–238
 - concessive conditional 226, 241–245
 - conditional 191, 226
 - core 251–252, 255–256
 - finite 182, 186, 213, 215, 217, 222, 227, 228, 236, 239, 240–244
 - illative 224, 236, 240
 - infinitive 182, 185, 186, 215, 216, 217, 227, 228, 236, 239, 244
 - main 201, 208, 215–218, 223, 228, 233, 236–237, 241
 - matrix 255
 - nominal 196
 - non-finite 186, 198, 204, 227,
 - relative (see: relative)
 - small clause 217
- clefting 252–254
- clitic 84–85, 93, 107, 112, 118, 133, 137–138, 154–155, 166, 175–176, 183, 192, 195, 197, 211, 217–218, 222, 251, 253–255, 262, 267, 275, 278, 390–391, 462, 623–624
- closing, closure 116, 376, 378, 381, 388, 475
- code 10, 29, 397–398, 413–414, 447, 511, 520, 604–605, 608–609, 611, 614, 685
 - code-choice 601 602 603 606 615 724 732 734
 - code conflict 511, 604
 - legitimate code 10, 605
- codeswitching 410, 605–615, 623
- codification criteria 82–83, 95, 96, 415
- coherence 288, 308
- cohesion 288
- collocations 271–289, 327
 - idiomatic vs non-idiomatic collocations 271–289
- Colón, Germà 52–54, 56, 61, 311–312, 315
- Comadira, Narcís 6–8, 591
- comarca 372, 374, 571, 579, 632, 638
- communication
 - inter-ethnolinguistic 604
 - interpersonal 604
 - participants (in the speech event) 133, 136, 139, 153, 188, 602, 606
- communicative event (see also: speech event) 399, 605
- community
 - communicative community 416–417
 - cultural community 29
 - fractal language community 638–640
 - language community 29, 414–417, 416–417, 441–442, 606, 619, 638
 - medium-sized language community 2, 645
 - national community 723, 726, 734
 - sociolinguistic community 414
 - speech community 275, 404–406, 410, 416, 617, 687, 694
- comparative construction 229–232
- competence
 - communicative competence 344, 401, 622
 - proficiency (see: language proficiency)
- composition 317, 365, 448, 466, 504, 525, 565, 569–570, 575
- compositional, compositionality
 - normative 9, 415–416, 687–688
 - semantics 272–278, 281–283
- compound
 - conjunctions 169
 - endocentric compound 281
 - exocentric compound 281, 365
 - forms 142, 158, 571
 - lexical compounds 115, 280
 - names 25, 569
 - nouns 365, 564, 566–568, 570, 574–576
 - place names 575
 - prepositions 182, 185
 - reduplicative compounds 367
 - relative pronoun 218–221
 - syntagmatic compound 280–281
 - tenses (see: tense)
 - word (see: word)

- compounding 103, 110, 329, 351–353, 363–368
- Compromise of Caspe 472
- computational techniques 422
- conceptual metaphors (see: metaphor(s))
- conceptual transparency/opacity 278, 283
- concessive constructions 241–243
- conditional(s) 144, 158–159, 162, 167, 174, 191, 198–199, 201–203, 208, 226, 241–245, 381, 465, 665
- conflate, conflated 110, 144, 148, 373, 385, 387
- conjugation
 - models 120, 140, 159–164, 373, 378, 380–382, 391, 475, 624
 - morpheme 159
- conjunction 131, 133, 135, 169, 185, 191, 206–207, 211–215, 217, 222, 224, 225–229, 231–232, 234–239, 241–244, 282–283, 303, 305, 385
- connective 201, 213–214, 238, 303, 305–308
 - discourse connective 213
 - parenthetical connective 135, 303–308, 608
 - pragmatic connective 303, 306–308
- connotation
 - negative connotation 323, 326, 392
 - pejorative connotation 361, 383
- conservative
 - dialects 155, 380, 391, 511, 567
 - politics 715
 - press 523, 559, 671
- consonant sequence (also: consonant(al) group) 381, 385, 390, 453
- consonant system 387
- consonant(al) inventory 389, 391
- consonants 102
 - affricates 87–89, 102–104, 113–119, 378, 387, 391, 406, 452–453
 - affrication 89, 113, 119, 387, 409
 - alternation 89, 96, 109, 117, 121–122, 149–150, 155, 162, 214–216, 291, 407, 449
 - alveolar 48, 103–105, 108, 113, 117, 407, 452–456
 - apicoalveolar 104, 105
 - approximants 102, 105, 117
 - assimilation 113, 116, 117, 381, 387, 449, 450, 451, 453, 454
 - bilabial 103, 104, 117, 388, 451
 - clusters 95, 103, 107–117, 120, 148–149, 151
 - continuant 88, 103
 - deaffrication 385, 406, 452, 453
 - dental 103–104, 112, 117, 382, 389, 452–453
 - devoicing 103, 109, 113, 116, 119, 383
 - dissimilation 115–112, 117, 379, 386–388, 450–451
 - elision 110, 118–119, 121
 - epenthesis, 110–114
 - fricative(s) 104–105, 109–110, 113, 118–119, 378, 387, 389–391, 406, 409, 451–455
 - glide(s) 104–114, 117, 120–121, 122, 146–148, 156, 161–162, 378, 448–449, 454
 - labio-dental 451
 - lateral(s) 90, 102–105, 108, 118, 386, 390, 454
 - lengthening 87, 104, 110–114
 - nasal(s) 89, 102, 104, 109, 114–117, 151, 376, 378, 385, 389–391, 456
 - obstruent(s) 88, 101, 102, 116–117
 - occlusive(s) 148, 151, 385, 390, 391, 453, 567
 - palatal 84, 103–105, 108, 114–115, 119, 146, 160, 374, 377, 378, 379, 387, 390, 408, 448, 454–455
 - postpalatal 103
 - prepalatal 104, 113, 117, 378, 389–391, 405–407, 455
 - rhotacism 565
 - rhotic(s) 90, 102, 105, 108, 111–113, 118, 382, 391
 - sonorant(s) 102, 120
 - sonority 102
 - sonority scale, 108–109, 120, 380
 - spirantisation 103–104, 117–118
 - split (depalatalisation) 110–114
 - stop(s) 95, 103, 111, 112, 113, 114, 117, 451
 - velar 103–105, 117, 160, 374, 379–382, 387, 391
 - voicedness 102–103, 109, 113, 116–117, 119
- Consorti per a la normalització lingüística* (CPNL) (also: Consortium for Language Normalization) 679, 700, 729, 730, 732
- constituent (see: syntactic constituent, word constituent)
- constraint 9, 83–85, 96, 272, 279–280, 298, 341, 377, 384, 517, 711
- content clauses 214, 223–224, 227
- context
 - cultural context 321, 324, 326, 335, 598
 - formal context 392, 731
 - phonetic context 377
 - pragmatic context 323

- sociocultural context 401, 410
- vocalic context 385, 390
- contextual value 83, 87
- contextualisation cues 248, 275, 606, 607
- contraction 85, 412, 572
- contrast 22, 103, 104, 119, 120, 451, 453, 454, 455
- contrastive 54, 206, 250–251, 253–254, 257, 266–267, 296, 300
- convention 83–85, 87–96, 274, 401, 518, 520, 525, 732
- conventionality 273–283
- Conventus Narbonensis* 438
- Conventus Tarraconensis* 438
- convergence 408, 490, 530, 559, 604
- conversion 351, 353–357, 365–367
- converted 148, 356, 359, 376
- coordination 211–213
- Coromines, Joan 6–7, 47, 51–52, 54–57, 77–78, 312–313, 379, 438, 453, 455–456, 550, 555, 557, 561–565, 572, 575
- corpus 9, 40, 49, 314, 320, 337, 414–415, 585, 701
- corpus, corpora (digital) 7, 253, 293–296, 306, 312–317, 334, 337, 386, 421–431, 697
- corpus planning (see also: language planning) 2, 49, 82–83, 413–414
- correlative structure 213
- Costura, la* 675
- Council of Constance (1414) 18
- Council of Europe 619, 650, 658–659
- counterfactual, counterfactuality 141, 242–243
- court ruling 671
- criterion 18, 82, 96, 365, 376, 380, 382, 417, 474, 551, 556, 688
- cultural scripts 321, 325
- Cyrillic 25
- Czech, Czechs 25, 34

- Darmesteter, Arsène 357
- Davis, Thomas Osborne, 17
- Decreto Romanones* (1902) 499
- defective 144, 152, 171
- definite article 84, 93, 107, 117, 132, 135, 138, 152–153, 170–173, 182–183, 186, 211, 220, 288, 294, 296–297, 308, 385, 387–388, 390–391, 458–460, 474, 568, 599, 623
- definite exclamatives 265
- deictic maintenance 294
- deictic neutralisation 294
- deictic shift 295
- deixis 171, 288, 291–298, 402
 - three-term space deixis 189, 291–292
 - two-term space deixis 167, 171, 189, 291–292
- demographic 474, 480, 485, 597, 618, 620, 630, 632, 634, 646, 653, 687
 - change 620, 632
 - minoritisation 618
 - structure 618
- demography of language, demolinguistics 478, 489, 629–646, 665–666, 675
- demonstrative 152–154, 170–172
 - adverb 173, 187–189, 240
 - determiner 172–173, 220, 288, 291–298
 - pronoun 154, 171, 180
- demonstrative(s)
 - distal vs proximal 171, 189, 220, 238, 288, 292–297, 386
 - non-situational vs situational 293–298
 - reinforced 292, 391, 459
- demonstrative system 153, 292
 - three-degree demonstrative system 153–154, 292, 385, 391
 - two-degree demonstrative system 153–154, 171, 291, 385, 391
- demonym(s) 19, 476, 574, 579
- Département des Pyrénées Orientales* 372, 549, 653–657,
- derivation 110, 118, 120, 329–331, 351, 353–354, 363, 365, 570, 575
- derivative nouns 564
- derive 187–188, 288, 299, 322–340, 357–364, 374–391, 443–467, 474–475, 563, 570, 573
- descriptive linguistics 38, 41, 44, 422
- descriptive phonetics 48, 53
- determinant 359–367
- determinatum 359–367
- determiner (see also: article; demonstrative; relative) 119, 132–137, 154, 170–172, 174, 218, 221–222, 255, 291–292, 459
- Deutscher Katalanistenverband* (DKV) 747
- diacritic 86, 93, 96
- diaereses 85, 87, 96–97
- dialect
 - boundary (see: boundary)
 - consecutive dialect 148, 373, 389

- constitutive dialect 148, 373, 474
- dialect(al) block 332, 373, 376, 379, 384, 511
- map 383
- region 382
- dialectal
 - differences 161, 371
 - diversity 83, 147, 152, 371, 390
 - division 373–380, 383–384
 - form 340, 392, 400
 - lexicon 332–336, 379
 - transition area 382–384, 406, 408
 - variation (see: variation)
- dialectological, dialectology 19, 38, 41, 43–44, 47, 50, 51, 53, 56, 58, 62, 378, 403, 406, 482, 529, 550, 699
- dialogism 612–613
- diaphasic variation (see: variation)
- diastratic variation (see: variation)
- diasystem, diasystematic, diasystematicity, diasystemic 82–83, 85, 93–97, 412, 415–417, 474, 687
- diatopic variation (see: variation)
- Diccionari Aguiló* 39, 47, 423, 524, 530
- Diccionari català-valencià-balear (DCVB)* 50, 77, 152, 312, 379, 423–424, 549, 535–536, 553, 555–556, 558, 562
- Diccionari etimològic i complementari de la llengua catalana (DECat)* 6, 55, 78, 312, 379, 557
- dictionary
 - of personal names 56
 - of pronunciation 415
 - of verbal usage 57
 - official dictionary 7, 47, 555
 - prescriptive dictionary 7, 56, 62
- differentiate(d) 85, 39, 54, 85, 272, 276, 337, 344, 380, 507, 534, 665, 689
- diglossia 338, 499–500, 518, 521, 536, 538, 546, 549, 581, 584, 631, 633, 642
- dignity 26–27, 491
- digraph (see: grapheme)
- diminutive 7, 115, 119, 130, 322, 328, 331–333, 352–353, 360–361, 454, 564, 569–575
- diphthongization 87, 448
- diphthong(s) 7, 106–108, 162, 448–449, 568
- disappear(ed) 104–106, 110–111, 151, 325–326, 328, 382, 387, 439, 488, 546, 568, 577–578, 583, 632
- disclaimer 271, 608
- discourse
 - discourse connective 213
 - discourse grammar 57
 - discourse markers 303, 306, 308, 610–611
- dislocation(s) 166, 183, 247, 249, 251, 252, 253, 254–257, 259, 265, 267, 275, 278
- distinctive, distinctiveness 83, 85, 93, 96, 97, 332, 456, 531
- distribution
 - functional 412, 479, 499, 501, 640, 642
 - geographic 333, 388, 408, 641
 - social variables 404, 408
 - structural 87, 89, 95, 142–143, 216, 276, 282, 288, 293, 296, 315, 332, 391
- disuse 38, 392, 530, 572
- documentation 319, 334, 379, 421, 426–428, 447, 455, 477, 488–489, 523, 530, 535, 548, 563, 574, 654, 696, 702–703, 739–740
- domain
 - conceptual domain 324
 - lexical domain 317, 337
 - semantic domain 314, 334, 340
- domain (see: Catalan linguistic domain)
- domains (see: domains of usage)
- dona* (as a discourse marker) 305–307
- Dorian, Nancy C. 609, 621–622
- double voice, double-voicing 605, 612–613
- dual route (in reading) 86
- duchies of Athens and Neopatra 472
- duplicate names 573
- durative 183, 229, 363
- Dutch 25, 324, 583, 591
- Ebre/Ebro (river) 30, 374, 437–438, 443–445, 455, 476, 478
- ecological 716
- education
 - adult language education 678, 729
 - Andorran education system 674
 - bilingual education 657, 669, 672
 - Catalan conjunction model 662, 671
 - compulsory education 520, 670, 673
 - environmental plan(s) 671
 - European Framework 671, 748
 - foreign language teaching/learning 59, 674, 680, 703, 748
 - L2 teaching 59, 63
 - language conjunction principle 677

- language education programmes 671–672, 673, 675
- language learner 44–45, 59, 679, 727
- language proficiency 632–641, 660–662, 670–673, 734, 748
- language(s) of instruction 654, 661–662, 670–672, 675–676, 678
- Eivissa 13, 292, 328, 333, 358, 372, 376, 387, 564, 638, 640, 662
- Eivissenc 13, 110–111, 113, 119, 153, 358, 383, 387
- Eivissenc-Formentera 13, 383, 387
- eixe* (*eix*), *eixa*, *eixos*, *eixes* (demonstrative) 153, 292, 385–386, 391
- Eiximenis, Francesc 28, 30, 480, 590
- el*, *lo*, *l'*, *'l*, *la*, *li*, *els*, *los*, *'ls*, *les* (third person weak pronoun) 107, 137–139, 155–157, 176–177, 390
- ell*, *ella*, *ells*, *elles* (third person strong pronoun) 133, 139
- el* (*lo*), *la*, *l'*, *els* (*los*), *les* (definite article) 8, 107, 152–153, 171, 182, 186, 220, 288, 385, 474
- el qual* (compound relative pronoun) 218, 220–221
- elative 361
- elocution 54, 58, 63, 540
- Els segadors* 510
- em*, *me*, *m'*, *'m* (first person singular weak pronoun) 137–139, 155–157, 176–177
- embodiment 279, 533
- emigration 482, 635, 644, 741–742
- emphatic 152, 191, 301, 389
- en*, *ne*, *n'*, *'n* (weak pronoun) 166, 176–179, 184, 390
- en*, *na*, *n'* (personal article) 153, 171, 288–289, 388
- enclisis
 - pronominal enclisis 462
- ending 84, 92, 111–122, 130–151, 376–389, 390, 391, 458, 464, 475, 512, 523–525, 573–575
- English 6, 11–12, 24, 27–28, 30, 32, 55, 85, 87, 106, 173, 185, 187, 205, 255, 263, 265, 281, 291–298, 301, 318–319, 332, 340, 343–344, 354, 358, 426, 429, 587, 601, 671–673, 675, 678, 701, 714, 735, 741
- Enlightenment 21–22, 487–488, 491–492
- ens*, *nos*, *'ns* (first person plural weak pronoun) 133, 139, 156–157, 159, 176–177
- es* (*ets*), *sa*, *s'*, *es* (*sos*), *ses* (definite article, *sa-lat*) 93, 117, 152–153, 171, 288–289, 387, 474
- es*, *se*, *s'*, *'s* (third person reflexive and reciprocal weak pronoun) 156–157, 177
- ésser* (auxiliary verb) 192, 465, 475
- estar* (verb) 109, 179, 181, 196–198, 205, 317
- este* (*est*), *esta*, *estos*, *estes* (demonstrative) 155, 292, 385, 391, 416
- Estonia 23, 27
- Estudis Universitaris Catalans* 525, 550, 676
- et*, *te*, *t'*, *'t* (second person singular weak pronoun) 156–157, 176, 461
- ethnic 9–10, 27, 29, 399, 404, 441, 445, 564, 566, 631
- ethnicity 443, 525
- ethnographic 4, 403, 602, 620, 723
- ethnography, ethnography of communication 397, 401, 403, 410–411, 620, 731
- ethnolinguistic
 - barrier 604
 - belonging 604
 - group 413, 602, 604
 - homogeneity 631–632
 - significance 603
 - vitality 713
- etymology 88, 95–96, 152, 527, 564, 568, 571
 - etymological form 390–391
 - etymon 95, 572, 578
 - false etymology 85
- Europe
 - European languages 40, 82, 423, 582
 - Mediaeval Europe 28
 - Western Europe 22, 28–29, 84, 336, 480, 486, 489, 503, 693
- European 2–3, 8, 10, 21–23, 28, 34, 40, 82, 85, 91–92, 371, 423, 430, 482, 486–488, 497, 503, 582–583, 591, 617, 619, 625, 631, 636, 649–650, 664, 671, 723, 729, 735, 742, 744, 746, 748
- European Charter for Regional or Minority Languages* 619, 650
- European Union (EU) 2, 21, 34, 601, 649, 664, 744
- exclamation 125, 249, 263
- experimental phonetics 41, 47–48, 50, 56
- expulsion of the Moriscos (1609–1614) 372, 486–487, 631

- extension (see also: increment) 140, 160, 219, 323, 333, 378–379, 474, 532
 – root extension 140, 160, 378–379
- Fabra, Pompeu 5, 7, 24, 38–40, 41–42, 45–49, 51, 57, 76–78, 82, 107, 146, 342, 352, 357, 376–379, 414, 522, 525–541, 546–548, 548–551, 556, 581, 601, 684, 710
- family origin 640
- Febrer i Cardona, Antoni 38, 491, 584
- Federació Internacional d'Associacions de Catalànística (FIAC) 739, 747
- Federació Internacional d'Entitats Catalanes (FIEC) 743
- Felibrige* 498, 601
- fer*
 – as a light verb 192–193
 – in impersonal temporal constructions 228–229
- Ferrater, Gabriel 40, 53, 61, 557, 591
- Ferrer, Joan 6, 379
- Ferrer, Vicent 22, 30, 32, 480
- feudalism 28, 314, 319, 330, 337, 445–447, 457, 461–462, 466, 571
- field of knowledge, area of specialisation 402, 693–694, 696–704, 705
- Fishman, Joshua A. 10, 402, 602, 616
- fluency (see also: competence) 598
- focus 247, 249, 250–254, 256, 258, 266
- footing 607, 711
- Forteza, Tomàs 39, 42, 44, 506
- France 6, 9, 18, 21, 24, 38, 51, 52, 54, 77, 318, 339, 371–374, 437–438, 472, 487, 498, 518, 549, 550, 565, 600–601, 608, 614–616, 621, 631, 633, 645, 649–652, 657, 676–677, 715, 726, 730, 740–741, 743, 747
- Franco, Francisco/Franco regime/Francoism/
 Francoist 51, 55, 77–78, 287, 316, 343, 541, 545–554, 556–558, 581, 587–588, 634–635, 643, 651, 670, 677, 678, 698, 710–711, 713, 726, 745–746
- Franja, la 9, 292, 374, 412, 616, 618, 646, 659–660, 674
- Frankish 314, 441–445, 449, 565, 573, 577
- Franks 440, 443, 445–446
- French 6, 18, 20–22, 28–31, 371, 382, 471, 488, 492, 499, 500, 507, 511, 519–520, 524, 530, 535, 538–540, 548–549, 572, 583, 591, 601, 611, 618–619, 631, 633, 651, 653–654, 657, 674–675, 715, 740, 742, 747
 – French Republic 549, 619, 651, 715
 – French Revolution 20, 488, 492, 618, 715
 – French State Council 657
 – government 21, 549
- French (language) 6, 12, 22, 28, 53–54, 59, 84–90, 92, 95, 105, 165–167, 221, 266, 300, 311, 316, 318, 322, 328, 335, 339–340, 342–344, 388–389, 401, 442, 481, 487–489, 491–492, 499, 511, 518, 520, 522, 572, 577, 581, 585, 590, 599, 607–608, 611, 616, 618, 621, 631, 633, 637, 641, 651–653, 657–658, 670, 675–676, 695, 700, 711, 735
 – French-Catalan interferences 54, 611
 – Quebec French/Quebécois 616, 695, 700
- function(s)
 – communicative 479, 605, 698
 – grammatical 130, 168, 352–353
 – language functions (also: functions of language) 616, 622, 624
 – social 97, 600, 603, 625, 697
 – syntactic 133, 138, 177, 214, 219, 222
- functional 312, 399–401
- functional variation (see: variation)
- funduqs (*alfòndecs*) 739
- fusion 145–146
- Fuster, Joan 53, 60, 78, 557–558, 588, 710
- Galicia 574, 741
- Galician (also: Galician-Portuguese) 24, 90, 442, 498
- Gallia Narbonensis* 439
- Gascon 89, 327
- Gaul 313, 437, 441, 445, 573
- gender-invariable, gender invariance 149–151
- Genoese 599
- genoelect 500
- genre 480, 482, 489, 499–500, 503, 512, 517, 586, 600, 602, 606, 621
- geolinguistic studies 56, 408
- geolinguistic variation (see: variation)
- German 21, 28, 31, 40, 42, 45–46, 509, 535
- German (language) 6, 11, 23–25, 30–32, 55, 59, 583, 587, 590–591, 631, 652, 663, 740, 747–748
- Germanic 25, 279, 281, 314, 358, 365, 440–443, 449, 565–566, 568, 572–573, 576–578, 747

- anthroponymy 565, 572–573, 576
- influence 442–443
- invasions 440
- peoples 25, 440–441
- toponymy 565–566
- gerund, gerund construction 158–160, 204, 227–228
- globalization 319, 328, 344, 650, 724, 734
- glossaries 482, 693, 696
- Gothic 28–29, 314, 565, 573
- Graeco-Latin grammar 40
- grammar
 - generative grammar 57, 403
 - historical grammar (see also: Old Catalan) 51, 55–56, 148, 312, 522, 555
- grammar shrinkage 491, 622
- grammatical content 145
- grammatical gender 137, 170
 - feminine gender 130, 147, 360, 376
 - masculine gender 147, 150, 376
- grammatical person (see also: person) 141, 145, 171
- grammatical property 136–141, 143, 198
- grammatical relationship 133
- grammatical specification 130, 144
- grammatical studies 41, 44, 55
- grammaticalisation 282, 320, 331, 352–353, 385,
- grammolect 500
- Gran enciclopèdia catalana* 383, 556, 699
- ‘grandmother factor’ 621
- grapheme 43, 83–84, 86–87, 90, 95, 409, 522–523
 - digraph/digraphs 86, 474
 - trigraph/trigraphs 86
- graphic
 - appearance 84, 96
 - convention (see: spelling convention) 84, 96
 - form (see also: written form(s)) 85
- grave and acute accents 38, 92
- Great Chronicles* 7, 480
- Greece 24, 27, 581–582
- Greek 6, 24, 26–27, 31, 331, 352, 362, 366, 437–438, 441, 481, 491, 564, 577, 584, 586, 599
- Grégoire, Abbé Henri 20
- ground 256, 266
- Guimerà, Àngel 81, 503–504, 508, 584, 590
- Gulsoy, Joseph 55–56, 61, 379
- Gumperz, John J. 10, 605–606
- gypsy, gypsies 608, 632
 - Catalan gypsies in France 9, 373, 399, 608, 611, 614–616, 620–621, 741
- Gypsy Catalan (see: Catalan) 614, 621
- Habsburg 474, 485–487, 489, 630–631
- hagiotoponyms 571–572
- Halliday, M. A. K. 400
- hanging topic left-dislocation (HTLD) 254–255
- Haugen, Einar 9
- haver* (perfect auxiliary verb) 199, 201, 203, 407, 409, 465, 475
- haver-hi* (verb) 195, 197–198
- Hebrew 6–8, 110, 479, 572, 599, 616
- Hellenist and Roman period 27
- Herder, Johann Gottfried 3, 21, 509
- heteroglossia 410, 612–613
- hi* (weak pronoun) 133, 155–156, 166, 177, 179–180
- Hispania 437, 439, 441–443
 - Hispania *Tarraconensis* 439, 441, 490
 - Hispanic kingdoms 19, 479
 - Hispanic monarchy 315, 486, 489, 493
 - Hispano-Visigothic 441
- historical language stratum 562
- historical morphology and dialectology 51
- historical phonetics 54, 62
- historical studies 56
- historical texts 55
- historiography of Catalan linguistics 61
- history of Catalan 4, 58, 318, 630
- history of Catalan linguistics 3, 37–67
- history of language science 44
- history of words 4, 63
- ho* (weak pronoun) 133, 139, 155–156, 166, 177, 180
- Hoenigswald, Henry M. 37
- Holy Roman Empire of the Germanic Nation 25
- home* (as a discourse marker) 305–307
- Homilies d'Organyà* 447, 476
- homography 86, 96
- homophony 86, 96
- humanist(s) 339, 480, 481, 482, 483, 582,
- Humboldt, Wilhelm von 3, 22, 63
- Hungarian 21, 23, 25, 31, 34
- Hungary 25, 741
- Hymes, Dell H. 411, 602
- hyphens 85, 87, 93, 96–97

- Iberian 47, 88–90, 311, 313–315, 317, 437–439, 443–444, 486, 488, 504, 549, 563, 576, 582, 599
- icon, iconic, iconicity, iconization 327, 615, 710, 713–715
- identity
 - act of 616
 - Catalan identity 508, 657, 713–714
 - collective identity 716
 - ethnic identity 9, 10
 - national identity 3, 18, 492, 724, 726–727
 - personal identity 12, 716
- ideology 3–4, 18–22, 31, 492, 509, 523, 598, 619, 633, 645, 670, 710, 712–713, 715–716
- idiom 104, 275–282, 625
 - idiom family 278, 279
 - idiomatic phrase 277, 280, 282–283
 - idiomatically combining expression 273, 277, 283
 - satellite-framed pattern 279
 - verb-framed pattern 279
- immigration 4, 19, 28, 313, 315, 316, 336, 553, 599, 603, 618, 632, 634–636, 638–640, 642–646, 651, 653, 663–664, 671, 674, 678–679, 711–712, 713, 723–735
- imperative 139, 141, 158, 160, 163, 203, 237, 249, 301
- imperfect indicative 144, 158–159, 161, 243, 381, 464–465
- imperfect subjunctive 144, 158, 160, 162, 203, 243, 382, 384, 389, 391
- imperfective 139–141, 146, 199, 203
- impersonal 144, 192–193, 197, 218, 228
- inchoative 140, 159–161, 378–379, 391, 416, 475, 623
- increment (see also: root extension) 379, 391
- indexicality 714, 716, 726, 732
- indicative mood (see: mood)
- Indo-European 40, 262, 313–314, 562–564, 599
- Indo-European languages 40, 262, 599
- industrial revolution 492, 617
- infinitive 118, 140–142, 158–159, 166, 170, 182, 185–186, 202, 204–205, 214–223, 227–228, 236, 239, 244, 354, 385, 463, 465–466
- infix 140, 160, 378
- inflected form 137, 140–156, 159, 161
- inflection 118, 354, 136, 142, 145–146
 - Inflection Phrase (IP) 251
 - nominal inflection 135–136, 136–139, 147, 151, 173, 281
 - verbal inflection 121, 135–136, 139–141, 146, 163, 198
- inflectional category (see also: grammatical property) 143
- inflectional marker(s) (see: marker)
- inflectional paradigm (see: paradigm)
- inflectional pattern 140, 159
- inflectional process 130
- inflective 130, 135–136, 142, 144, 150, 191, 565
- information
 - given information 227, 248
 - grammatical information 145–146
 - information focus (see: focus)
 - information packaging 248, 256
 - information structure 255–256, 266
 - new information 197, 250, 256
- Iniciativa Cultural de la Franja 678
- Institut d'Estudis Catalans (IEC) 2–3, 11, 24, 39–40, 45, 61, 77, 81, 316, 319, 342, 351, 413, 415, 424, 428, 504, 512, 517, 527, 538, 546, 552, 601, 604, 657, 676, 698, 714, 746–747
- Institut Ramon Llull (IRL) 590, 680, 739, 748
- integration 317, 365, 477, 532, 553, 603, 646, 670, 672, 711, 723–724, 728–729, 731–732, 734
- interface 258, 263, 423, 425, 427–430
- intergenerational transmission 550, 618, 631, 645–646
- interjection 131, 133, 135, 153, 299
- internal temporality 140
- international normalisation 696, 705
- interrogative sentence 207, 249, 258, 260–261, 263
- interrogative words 258, 259–260, 262–263
- intonation
 - declarative sentences 122–124
 - exclamations 125–126, 263–266
 - focused information 123
 - non-polar questions 125
 - polar questions 124
- intonational phrase 255–256
- invited inference 323
- Ireland 17, 18, 24
- irregular verbs (see also: verb) 44, 146, 158
- isogloss 333, 373–374, 378, 382–384, 455
- isolationism 414
- isomorphism 327

- Italian 6, 12, 21–22, 28, 84, 86, 92, 165–166, 266, 315–316, 318, 322, 328, 335–336, 340, 342–344, 373, 381, 401, 438, 481–482, 498, 511, 522, 567, 577, 582–584, 587, 590–591, 599, 602, 607–610, 619–620, 622, 633, 637, 641, 652, 655–656, 675, 739, 745–746
- Italic 440, 442, 479, 608, 740
- Italy, Italian Republic 38, 335, 339, 371, 439–440, 473, 476, 481–482, 498, 581, 599, 619, 633, 649–650, 652, 677, 726, 740
- iterative 204
- Jespersen, Otto 44
- Jewish 479, 571, 583, 599
- jo, mi* (first person singular strong pronoun) 133, 138, 139, 155, 175–176, 461
- Jocs florals* 49, 503, 507, 508, 511, 519–521, 533, 541, 547, 552, 742, 745
- Labèrnia, Pere 39, 491, 518–519
- Labov, William 397, 402–405, 410
- language
- attrition 405, 408, 598, 606, 622–624
 - awareness 501
 - change (also: linguistic change) 83, 85, 335, 388, 406, 631, 633
 - choice 601–603, 606, 615, 724, 732, 734
 - community (see: community)
 - conflict 23–24, 488, 602, 604, 633
 - contact 42, 60, 105, 314–319, 325, 328, 331, 335–336, 340, 375–377, 389–390, 408, 475, 490, 597–626, 630–631
 - correctness 53, 689
 - death 12, 598, 616, 622
 - decay 12
 - domain, linguistic domain (area) 335–336, 372, 570, 579, 625, 690
 - domains (domains of language use) 336–342, 480, 504, 602, 617, 619, 622, 630, 640–641, 645–646, 654, 661–662, 676, 711–712
 - dominant language 411, 413, 520, 598, 617, 619, 621–622, 638, 641, 716
 - ecology 597–599
 - erosion 408, 714
 - established language 82, 97, 411
 - first language 412–413, 581–582, 623, 631, 639–641, 643–644, 670–672, 713, 715
 - formal language 25, 153, 184–185, 300, 398, 557, 630
 - function 616, 622, 624–625
 - history 9, 58, 63, 312, 337, 629
 - ideologies
 - anonymity 709–712, 716
 - authenticity 11, 24, 709–712, 714, 716
 - diglossic ideologies 711
 - “ecological” ideology 716
 - erasure 713–714, 726
 - fractal recursivity 713–715
 - fragmentation 715
 - *gonellisme* 715
 - Herderian ideology of language 3
 - Humboldtian ideology of language 3
 - iconization 713–714
 - integrationist 670
 - *llemosinisme* 493
 - *mossarabisme* 535, 538, 559
 - national language ideology 3, 18, 20, 22, 633, 635, 645, 724
 - recursive particularism 714
 - institutionalization (also: institutionalisation) 457, 517, 519, 528, 536, 546
 - interference 59, 328, 331, 340, 408, 410, 411, 412, 413, 415, 490, 611, 622
 - intergenerational transmission 408, 412, 550, 618, 631, 633, 645–646
 - Language Normalization Act 78, 673
 - ‘language pain’ 710
 - legitimate language 604
 - literary language 5–8, 25–33, 48–49, 82, 454, 475, 480, 512, 526–529, 534, 584, 601
 - loyalty 598, 602, 616
 - maintenance 617–619
 - major language 2, 507
 - minoritized language 397
 - minority language 1, 3, 479, 488, 598, 619, 650, 652, 655–656, 698
 - models 26, 33–34, 335, 340, 398, 402, 582, 675, 686, 691
 - national language (see also: Castilian, Catalan, French, Italian, Spanish) 20, 27, 488, 492, 519, 587, 724, 735
 - normalisation (also: normalization) 25, 78, 316, 412, 602, 650, 660–670, 673, 679, 700
 - obsolescence, obsolescent language 611, 622–624
 - of culture 25, 39, 478, 581–582, 599, 631, 633, 702

- official language (see also: Castilian, Catalan, Spanish) 2, 24, 27, 31, 60, 633–634, 649, 664, 674, 676, 730
 - oral language (see also: spoken language) 42, 416, 490, 600, 615, 688–689
 - planning 3, 10, 344
 - corpus language planning 49, 82–83, 413–414, 417
 - status language planning 9, 545, 630, 655, 700
 - policy/policies (see also: education) 8, 60, 414, 498, 549–550, 597, 635, 643, 645, 649–666, 670, 672, 676–678, 699–700, 714, 724, 727, 729
 - purified language 24
 - receding language (also: recessive language) 12, 408, 486, 598, 617, 619, 621–622
 - reduction 622
 - reform 5–8, 39, 41, 51, 82, 93, 96, 97, 107, 446, 517–541, 710
 - regional language 20, 652, 657, 676
 - renaissance 23, 48, 519
 - repertoire (see: repertoire)
 - representations 82, 96
 - resources (see: resources)
 - restriction 339, 401, 404, 411–412, 530, 541, 598, 622
 - revitalisation (also: revitalization) 10, 618–619, 636
 - revival 5, 78, 497, 498, 503, 504, 555, 589
 - shift 10, 314, 412, 518, 597–598, 602, 610, 614, 616–622, 632–634, 638, 640, 642–646, 678, 715
 - reversing language shift 10, 116, 602, 618
 - shrinkage 408, 491, 622, 624
 - small language 2
 - spoken language (see also: oral language) 42, 155, 150, 158, 300, 380, 416, 446, 490, 497, 512, 522, 546, 600, 615, 685, 688–689
 - standard language (see also: variety, varieties) 25, 49, 59, 203, 239, 288, 329, 334, 458, 517, 522, 555, 602, 687, 688, 689, 691,
 - state language(s) 23, 34, 631, 633, 635, 642, 646, 649
 - status 414
 - strata 313, 562, 567
 - target language 582
 - teaching 59, 638, 678–679, 729
 - unified language 24
 - usage 59, 63, 340, 499, 553, 621, 650, 665
 - use 9, 392, 397, 402, 410, 499, 546, 602, 604–605, 615–616, 629, 640–641, 656, 677
 - variation (see: variation)
 - varieties (see also: Castilian; Catalan; Occitan; Romance; Valencian; variety, varieties) 288, 605
 - written language 17
- Latin
- Classical Latin 376, 463–464
 - diversification of spoken Latin 440
 - Greco-Latin vocabulary 338
 - Hispanic Latin 313, 438
 - inchoative verbs 140, 160
 - influence of written Latin 450
 - Latin alphabet 84
 - Latin declensional system 458
 - Latin fourth conjugation 381
 - Latin idiom 104
 - Latinate, Latinized 7, 329–330, 359, 447
 - Latinising style 481, 564–565
 - Latinization 441
 - popular Latin 437
 - vocabulary 337
 - Vulgar Latin 4, 149, 155, 371, 373, 376–377, 390, 448, 451, 463, 465, 474–475, 490, 562
 - writing 338, 354
 - written 440, 447, 630
- L'Avenç* 39, 49, 76, 504, 512, 522–525, 533, 585
- Law of Regional Language Policy 676
- Law of Use and Teaching of Valencian 671
- Law on Plurilingualism in the Educational System 673
- learned helplessness 603–604
- L1/L2 learner 44–45
- legislation 337, 499–500, 520, 545, 550, 554, 559, 618, 619, 651–654, 657–658, 660, 662–663, 676
- legitimacy 10–11, 614–615, 710
- lemmatization 424
- level of formality 400–401
- levelling 412, 442, 474, 622
- lexeme 187, 353, 360, 365–366, 490, 568, 574–575

- lexical category (see also: word classes) 131, 134
- lexical expression 282
- lexical hispanization of Catalan 475
- lexical reform 51
- lexical syntax 272
- lexicalized phrase(s) 271, 273, 275, 280–283
- lexicographical, lexicography 39, 41, 43, 46, 48–49, 50–51, 52, 55, 56, 58, 62–63, 352, 423, 491, 518, 524, 529, 685, 698–699
- lexicology 58, 320
- lexicon
 - general lexicon 2, 311–346
 - mental lexicon 272, 279, 283
- Ley de Enjuiciamiento Civil* (1881) 499
- Ley del Notariado* (1862) 499
- Liber Iudiciorum* 447
- light verb construction 272
- Limousin 493, 501, 511
- lingua franca* 413, 476, 728, 732, 735
- linguistic analysis 38, 312, 422
- linguistic attitudes (also: language attitudes) 10, 343, 414, 417, 645, 712, 713
- linguistic authority 709–711, 715
- linguistic censuses 633
- linguistic community (see: community)
- linguistic conservatism 387, 475, 518, 714
- linguistic content 145
- linguistic cosmopolitanism 716
- linguistic differentiation 712
- linguistic minority 652
- linguistic normalisation principle 677
- linguistic particularism 554
- linguistic repertoire (see: repertoire)
- linguistic research 45, 63, 273
- linguistic resources (see: resources)
- linguistic restriction(s) 106, 107, 178, 222, 236, 238, 273, 277, 327
- linguistic rights (also: language rights) 8, 478, 536, 616, 649, 651–652, 657, 660–665
- linguistic studies 41, 43, 52
- linguistic sustainability 716
- linguistic theory 320, 422
- linguistic transparency/security principle 678
- linguistic variation 56, 398–402, 403–404, 410, 477, 695
- linguistic works 44, 50, 62
- linguistically mixed families 713
- linguistics
 - applied linguistics 50, 58–59, 63, 703
 - cognitive linguistics 278
 - computational linguistics 58
 - contemporary linguistics 44, 55, 61
 - contrastive linguistics 54
 - corpus linguistics 403, 422
 - general linguistics 53, 402
 - geolinguistics 63, 404, 406,
 - historical linguistics 38, 39, 44, 55, 404, 568
 - Romance linguistics 1, 38, 40, 45–46, 51–53, 56
 - text linguistics 287, 308, 402
- link 256–257
- literacy 8, 32, 489, 492, 633
- literary
 - language (see: language)
 - ‘literary independence’ 34
 - renaissance 48
 - use 480, 503, 511
- Lithuania, Lithuanian 23, 34
- llemosí* (see also: Limousin, Provençal) 476, 493
- llemosinisme* (see: language ideologies)
- llengua comuna* (‘common language’) 600, 687, 711–712, 724, 728, 734
- llengua pròpia* 660, 663, 710–711, 724, 727–728
- Llull, Ramon 7, 29, 358, 423, 427, 478, 480, 590–591, 655, 663, 739, 744
- llur, llurs* (third person plural stressed possessive) 139, 155, 389, 460
- Lo Gaiter del Llobregat* (see: Rubió i Ors, Joaquim) 34
- loanword 7, 48, 83, 85, 103, 105, 608
- localisation 362
- locution 135, 182, 185–188, 212–213, 221, 228, 231, 239–244, 248–249, 301
- logogram 85–86, 93
- Loi Deixonne* 549–550, 657
- Macedonia 22, 31
- Magyarisation 25
- Majorca
 - Kingdom of 19, 28, 472
- Majorcan 13, 56, 92, 103, 107–118, 120, 375, 383–384, 387, 521
- majorités fragiles* 723
- Mansi, Joannes Dominicus 18
- Manuscripts 456

- late-Latin manuscript 446
- Marfany, Lluís 669
- markedness 142–143
 - marked forms 142, 143, 622
 - unmarked forms 6, 110, 142, 143, 206, 223, 326, 337, 623
- marker
 - absence of markers 142, 145
 - class marker 158–159
 - conditional marker 162
 - conjugation(al) marker 159
 - dative marker 139, 183
 - feminine marker 130, 142, 147–150, 162
 - future marker 145, 162
 - gender marker 145–152
 - gerund marker 159
 - imperfect indicative marker 161
 - imperfect subjunctive marker 382, 384
 - inchoative marker 160, 161, 378–379, 391, 416, 475
 - inflectional marker(s) 130, 135, 142–147, 150, 152, 158–163
 - number marker 145–150
 - past indicative marker 161
 - person marker 143, 158–163
 - plural marker 117, 130, 139, 142–151
 - present subjunctive marker 162
 - tense marker 143, 146, 158–163
- mass media
 - broadcasting media 539–540, 597, 654–655, 661–663, 683, 690, 744,
 - language advisors 604, 683–686
 - oral media 59
 - print media 683–690
- Massó i Torrents, Jaume 39, 76, 521, 523, 525
- Maugham, William Somerset 62
- meaning
 - amelioration of 323
 - broadening 322
 - compositional vs non-compositional 272–284
 - context-bounded 613
 - conventional 273–275
 - deictic 132
 - emotive 323, 360
 - epistemic 205–206
 - figurative 274
 - grammatical 135
 - idiomatic 275–282
 - intended, intentional 305, 606
 - interactional 298, 306
 - intersubjectification 324
 - lexical 130, 185, 352, 353, 362, 364
 - lexicalized 281
 - literal 274, 277, 282
 - narrowing 322
 - original meaning 323, 328, 375
 - pejoration of 323–324
 - phrasal 275, 328,
 - pragmatic 237, 288, 298, 308, 608
 - propositional 198, 206, 227, 237, 248, 288, 298, 608
 - referential 306, 400
 - social 605, 712
 - subjectification 324, 331
 - textual 288
 - transparent 278, 283
- Mediterranean 29, 311, 313, 315, 318, 371, 437, 439, 453, 472, 476, 479, 481, 486, 564, 599, 614, 739–740
- Meillet, Antoine 18, 27
- Mel'čuk, Igor 271
- Menéndez Pidal, Ramón 44, 51, 524–525, 532
- merge 102, 117, 146, 451–453
- metalinguistic 493, 609
- metaphor(s), 4, 21–22, 312, 321, 339, 481, 480, 533, 538, 549, 582, 586
 - cognitive mechanism 4, 324
 - conceptual metaphors 278–279, 283, 330
 - semantic change 324–325
- metaphoric(al)
 - metaphorical codeswitching 606, 608
 - metaphorical motivation 283
 - metaphorical nature of human thought 279
- metapragmatic 609, 613
- Metge, Bernat 44–45
- methods
 - corpus analysis 291, 311, 327
 - diachronic 333–335
 - diachronic sociolinguistics 311, 337
 - empirical 422
 - ethnographic 620
 - historical-comparative 39
 - language history 337
 - qualitative 334, 410, 620
 - quantitative 404, 411, 620, 703
 - sociolinguistic interview 404, 409, 411
 - sociological 389, 602, 620
 - statistical 411

- metonymy 4, 324–325, 330
 - cognitive mechanism 4, 324
 - metonyms 365
 - semantic change 324–325, 330
- meu, meva/meua, meus, meves/meues* (first person singular stressed possessive) 154–155, 460
- Mexico 31, 33, 547, 742–744, 747
- Meyer-Lübke, Wilhelm 44, 745
- Middle Age(s) 312, 331
- migration 10, 331, 413, 539, 597, 606, 611, 617, 622, 632, 634, 642, 670, 716, 723–730, 734, 741–743
- Milà i Fontanals, Manuel 38, 374, 511–512, 521
- Minorcan 13, 107, 110–112, 114, 117–118, 383, 387
- minority, minorities 1, 3, 9, 12, 20, 25, 49, 148, 175, 373, 382, 390, 410, 414, 479, 486, 488, 492–493, 507, 551, 598–599, 617, 619, 630–631, 641, 650, 652, 655–656, 698, 705, 723
- modalisation (also: modalization) 298, 308, 402
- modality 198, 214, 242, 247–248, 266–267, 288
 - declarative modality 122, 125, 168, 206, 214–215, 278, 302
 - exclamatory modality (also: exclamative modality) 250, 263
 - interrogative modality 226, 258
- Modern Age 312, 315, 316, 317, 327, 331, 339, 517, 519, 527
- Modern Catalan 5–6, 42, 150, 157, 320, 379, 458, 530, 537, 601
- Modernisme* 584, 586
- modification 277–278, 352, 359–363
- Moll, Francesc de Borja 43–44, 46, 50–53, 56, 59, 61, 78, 379, 450, 532
- mon, ma, mos, mes* (first person singular unstressed possessive) 139, 154, 171–172, 460–461
- monolingual(s) 499, 581, 603, 631, 633, 642, 646, 665–666
- monolingualism 479, 552, 715, 716
- monomorphemic 130
- monophthongization 448
- Montoliu, Manuel de 39, 46–47, 423, 537, 548
- Monzie, Anatole de 20
- mood 129, 136, 139, 141, 146, 158, 223, 226, 555
 - imperative mood (see also: imperative) 139, 141, 163, 198
 - indicative mood (see also: indicative) 139, 141, 146, 167, 198, 216, 223, 625
 - subjunctive mood (see also: subjunctive) 120, 139, 141, 146, 147, 223–227, 298, 379
- Moorish 444–446, 486
- Moreu-Rey, Enric 56, 562, 574
- Moriscos* 372, 486–487, 631
- morpheme 83–84, 86, 95, 103, 105, 114, 117–119, 121–122, 129–130, 145, 150, 159, 161, 166, 354, 366, 456–457, 465, 623
- morphogram 84, 86
- morpholexical typology 574
- morphological, morphology 4, 39, 44, 50–51, 56–57, 61, 63, 83–86, 93–97, 111, 113–115, 120, 129–131, 136, 145, 147, 149, 158–159, 161, 248, 251, 272, 318, 329–330, 334, 351–354, 361, 365, 376, 380–381, 387, 389, 415–416, 442, 449, 501–502, 528, 531, 562, 574, 623–624, 687
- morphological formation 562
- morphological segmentation 130, 147, 159, 161
- morphological transparency 84–85, 95–97, morphophonemics 118–120
- morphophonology, morphophonological 177
- morphosyntactic features of early Old Catalan 457–467
- morphosyntactically annotated corpus (see also: corpus, corpora (digital)) 427–429
- morphosyntax 53, 57, 63, 129, 537
- Mozarabic (see also: *Andalusi* Romance) 314–315, 374, 446, 475, 567, 570, 573
- Mozarabic liturgical rite 445
- ‘Mozarabism’ (*mossarabisme*, see: language ideologies)
- Mudejar 478, 599
- multiculturalism 725
- multilingualism 426, 598, 601, 640, 645, 704, 716, 724
- multi-word expression 272
- Murcia 372, 374, 438, 616, 634, 651, 658
- Murcia, Kingdom of 472, 632
- Muslim 19, 26, 314–315, 372, 446, 473, 566, 631–632
- Musset, Alfred de 20–21

- mute consonant 148, 150
- mutual recognition 680, 716
- names
 - compound names 569
 - duplicate names 573
 - first names 288, 546, 578
 - forename(s) 576–578
 - name of the saint 572
 - nickname(s) 6, 562, 576–578
 - personal names 12, 56, 170, 288, 308, 313, 442–446, 457, 561–577, 660
 - place names 12, 182, 375, 388, 443, 446, 546, 561–576, 656–662, 740–741
 - proper names 12, 341
 - surname(s) 170, 442, 562, 576–579, 655, 741
- Naples, Kingdom of 472–473, 479, 481–482, 740–741
- narrative 305, 410, 481, 607–610, 620, 630, 725, 726
 - sequences 295
 - texts 289, 401, 426
- nation
 - nation building 506, 510
 - nation-state 3, 12, 602, 617, 724–725
- national identity (see: identity)
- national ideology 18, 22, 619
- national question (see: question)
- nationalism 21, 24–25, 505–506, 523–525, 531, 630
- nationality/ies 17, 18, 23, 28, 509, 525–526, 536, 631, 635, 654, 724, 726, 729
- Neapolitan 476, 479, 599
- negation 167, 206–208, 225–226, 232, 237, 245, 262, 299–300, 302, 362
- negative polarity 206–207, 300
 - negative polarity item 167, 207
- Neogrammarian 39, 44
- neology 3, 58, 63, 696, 704
 - neologisms 3, 338, 340, 344, 366, 424, 429, 700, 704
- Neology Observatory (OBNEO) 704
- neutralise(d) (also: neutralize(d)) 162, 296, 329, 399, 491
- NGO 724, 728–732, 734–735
- Ninyoles, Rafael Lluís 60, 603, 715
- no (negative adverb) 166–167, 206–208, 232, 299, 300, 389
- nominalisation 354, 363
- nominative 131, 138, 155, 255, 458, 461
- non-assertive 141, 198, 223
- non-Indo-European 313, 562–563
- non-inflective 130, 135
- norm
 - monocentric norm 398, 414–415
 - pluricentric norm 398, 414
- Normes de Castelló* 531, 554, 685
- North American Catalan Society (NACS) 746–747
- Northern Africa 313, 443, 744
- Norwegian 9, 24
- nosaltres* (first person plural strong pronoun) 139, 156, 175–177, 183, 291
- nostre, nostra, nostres* (first person plural stressed possessive) 134, 139, 155, 172, 291, 460
- Noucentisme* 40, 76, 534, 557, 586, 600
- noun
 - agent noun 353
 - collective noun 173, 175, 353, 360, 364
 - common noun 170
 - countable noun 137, 170
 - measure noun 174
 - noun composition 575
 - place noun 353
 - proper noun 170–171, 288, 290–291, 385
 - simple noun(s) 566, 568, 570–571, 574
 - uncountable noun 131, 170, 174
- noun phrase (NP) 131–132, 137, 168, 170, 172, 293, 298, 353
- nucleus 106–107, 131–132, 230–232, 234
- number 136–138, 139–140, 136–163
 - plural number (see also: plural) 130, 137
- numeral 173, 175, 181
- object
 - cognate object 194
 - direct object 138, 169, 177–183, 192–194, 215–223, 247, 251, 259, 274–280, 367, 462, 466
 - indirect object 79, 138, 169, 177, 179, 183, 193, 215, 217, 223, 247, 255, 259, 462
- Obra Cultural Balear 53, 78, 678
- Obra Cultural de l'Alguer 678
- oc* (Occitan form) 474, 611
- Occitan (language) 7, 84, 88–92, 165–167, 300, 311, 314, 317–318, 322, 331–335, 372, 381, 388, 439, 474–481, 493, 539, 541, 567, 573, 577, 590, 599, 671

- variety 25, 327, 442, 448, 459, 460, 471, 493
- vocabulary 311, 316–317
- Occitania 311, 315, 471, 474
- officiality principle 677
- oi* (Catalan discourse marker) 610–611
- Old Catalan
 - analytical perfect paradigms 465–466
 - development of consonants 450–457
 - development of vowels in Old Catalan 447–450
 - morpho-syntactic developments in Old Catalan 457–467
- Old Catalan documents 458–465
 - *capbreus* 446
 - *clams* ‘complaints’ 447, 461
 - *querimònies* ‘accusations’ 447, 461
 - *rancures* ‘resentments’ 447, 461
- Òmnium Cultural 556, 678
- Onomasticon Cataloniae* 55, 78, 562
- onomatopoeia 342, 367
- opposition
 - binary opposition 136, 140, 185
 - gender opposition (also: masculine/feminine opposition) 136, 142, 149
 - markedness opposition 206, 223
 - number opposition (also: singular/plural dichotomy) 136, 141, 143
 - phonological opposition 102, 105
 - three terms opposition 143
- oral standardisation 63
- orality 8, 338, 613
- orthographic, orthography (see also: spelling) 82, 83, 85, 86, 96, 97
 - alphabetic orthographies 83–84, 93
 - orthographic norm(s) 11, 82, 391, 414–415, 505
- orthology 54, 58, 63
- Oscan 440
- overmarking 294–298,
- own language (see also: *llengua pròpia*) 554, 660–665, 677, 690, 710–711, 724–728
- oxytone (see: stress)
- palatalization of word-initial L 90, 456
- Pallarese 384
- Par, Anfós 44–45 533, 585
- paradigm
 - inflectional paradigm 142, 622
 - nominal paradigm 142
 - verbal paradigm 142–144, 158, 380–381
- parasynthesis 357
- parenthetical connective (see: connective)
- Paris 31, 47, 51, 53, 54, 77, 547, 744, 746
- PAROLE Catalan Corpus 429
- paroxytone (see: stress)
- participle 7, 42, 136–138, 204, 217, 322, 330, 353, 354–355, 363, 367, 465–466
- participle construction 227, 228, 234, 236
- particularism 414, 492, 554, 714, 715
- partitive construction 175, 174–175, 184
- pas*
 - modal particle 298–302, 308
 - negative adverb 166–167, 206–207, 232, 299, 389
- passive construction 137–138, 192, 276
- past
 - narrative past 202
 - periphrastic past (see: tense)
 - simple past (see: tense)
- patient 194, 277
- patriot, patriotic, patriotism 17, 21, 42, 503, 510, 529
- patronymic, patronyms 566, 578
- people
 - people’s personality 3
 - “peoples without history” 22
- perfect
 - analytical perfect paradigms 465
- perfective 139, 140, 199, 200, 203
- Perpinyà/Perpignan 13, 30, 43, 373, 375, 399, 520, 524, 527, 530, 535, 565, 614, 620, 657, 676
- person
 - first person 7, 138, 139, 141, 143, 146, 154–157, 160–162, 163, 175, 182, 250, 373, 379, 380–382, 385, 389, 391, 475, 728
 - second person 139, 141, 143, 144, 146, 154–155, 163, 166, 175, 385, 475
 - third person 136, 137, 138, 139, 140, 141, 143, 144, 145–147, 154–156, 161, 163, 175, 198, 290–291, 366, 378, 380–382, 384–386, 389, 391, 460, 625
- personal names (see: names)
- personal pronoun (see: pronoun)
- personal verb form 300, 407, 466, 625
- Pessoa, Fernando 21
- Petit i Aguilar, Joan 38, 518

- phoneme 83–88, 92, 385–387, 391, 451–456, 523
- phonemic, phonemic principle 83–96, 385
- phonetic distinction 149
- phonetic law 147
- phonetic realization 152, 388
- phonetic transcription 56, 62
- phonetics 4, 41, 44, 45, 47–48, 50, 53, 54, 56–57, 58, 61, 62, 63, 101–110, 687
- phonological, phonology 101–126
 - content 145, 150, 161
 - form 145–146
 - opposition 102, 105
 - transparency 84
- phoric marker 291, 297
- phrase
 - adjective phrase 168, 180, 271
 - adverb phrase 168, 187, 186
 - idiomatic phrase 277, 280–283
 - intermediate phrase 255–256
 - noun phrase 131, 132, 137, 168, 172, 293–298, 353
 - preposition(al) phrase 282
 - verb(al) phrase 168, 206, 273–275, 283, 385
- phraseme 271–284
- phraseology 39, 271–273, 317, 322, 325, 339, 365, 402
- Pityusic Islands 13, 372, 386, 472, 476, 532
- pla or romanç* 19
- place names (see: names)
- plural 117–120, 136–138, 139–140, 136–163
- pluricultural 671
- plurilingual 671–672
- plurilingualism 417, 479, 661, 672, 715
- point text maps 378
- polarity 167, 188, 206, 207, 300, 307
- policies
 - assimilation policies 650, 666, 724
 - autonomous policies 651
 - education(al) policies 603, 674
 - language policies 414, 549–550, 635, 643–651, 660–666, 670, 714, 724
 - language-in-education policies 669, 671, 673
 - public policies 633, 413
 - top-down policies 10, 602, 619
- political persecution 485–487
- politics 3, 4, 9, 26, 27, 31, 340, 488, 538, 665, 731, 732
- polymorphemic 130
- polymorphic 9, 415, 416, 687, 688
- polyphony 613
- polysemy 327
- Portugal 19, 27, 482, 487, 677
- Portuguese (language) 12, 21, 24, 84–92, 117, 166, 266, 314, 322, 590–591
- possessive 131–139, 147, 154–155, 170–172, 177, 186, 221–222, 280, 291, 389–391, 459–460
- posteriority 140, 228
- postnominal 131, 134, 172, 180, 194
- postposed 131
- power 8, 9, 12, 19, 23, 26, 30, 31, 315, 440, 445, 472, 473 486, 488, 498, 503, 508, 519, 527, 598–600, 605, 635, 664, 645, 672
- power and solidarity 6, 306, 602
- practices
 - communicative 614, 636
 - language/linguistic 492, 614, 631, 636, 641–642, 717, 724, 725, 732
 - local 602
 - social 30, 598, 602, 608, 614, 724, 725
 - religious 598, 599
 - writing 502, 616, 725
- pragmatic connective (see: connective) 303–306
- pragmatics 59, 63, 123, 263, 287–308, 397, 399, 402, 404, 410–411
- Prague 25, 590
- Prague Linguistic Circle 5, 44, 82
- Prat de la Riba, Enric 3, 40–46, 63, 77, 81, 508–528, 601
- pre-Catalan Romance 567
- pre-Christian 571
- predicate 132, 140, 166–168, 178–179, 184, 188–196, 214, 217, 226, 236, 239–240, 277, 353–355, 362
 - nominal predicate (see also: predicate nominative) 137–138
 - predicate nominative 131
 - secondary predicate 131
- prefix, prefixation 96, 109, 351–365
- prenominal 131, 134, 172, 180–181
- pre-Roman (see: Rome)
- preposition 181–188, 215, 220–221
- preposition phrase (PP) 168, 181–183
- prepositional locution 182–186, 221, 239–244

- prepositional regime supplement 169, 182–184, 193
- prescriptive corpus 9, 49, 414–415, 685
- prescriptive dictionary (see also: dictionary) 7, 48, 51, 56, 62, 82, 697, 705
- prescriptive grammar 48, 57, 77, 82, 198, 518, 535
- prescriptive guidelines 89, 684–691
- prescriptive issues 351
- prescriptive norms 8, 61, 687
- prescriptive rules 522
- prescriptive syntax 57, 714
- present
 - indicative 106, 140–146, 158, 159, 160, 161, 162, 163, 243, 379–391, 463–465, 475,
 - perfect indicative 158, 466
 - subjunctive 113, 144, 158–162, 203, 243, 379, 381, 389, 391, 466
- prestige
 - social prestige 479–480
 - sociolinguistic prestige 597–598
- presupposition 206, 288–308
- Primer Congrés Internacional de la Llengua Catalana* (also: ‘First International Congress on the Catalan Language’) 40–49, 504, 704, 745
- productive 135, 187, 329, 442, 565
- progressive (political) 505, 520, 521, 525, 530, 672, 742
- pronominal 107, 135–136, 166, 180, 194, 195, 276, 386, 462, 475
- pronoun
 - clitic object pronoun (see also: unstressed pronoun) 175
 - demonstrative pronoun (see: demonstrative)
 - interrogative pronoun 234, 259, 462,
 - personal pronoun 133–139, 147, 155–156, 186, 197, 291, 461
 - relative pronoun (see also: relative) 183, 218–222
 - resumptive pronoun 254–255
 - stressed personal pronoun 139, 461
 - strong pronoun 155, 175–177, 255
 - unstressed pronoun (see also: weak pronoun) 462
 - weak pronoun 155–157, 165–166, 176–193, 255, 291, 298, 385, 475, 714
- pronounced 7, 85–95, 103–120, 144–164, 366–389, 456–457
- pronunciation 43, 54, 63, 85, 91–117, 148–149, 373–376, 381, 387–389, 409, 415, 455, 522, 530, 531, 538, 613, 683–684
- proparoxyton (see: stress)
- prosodic 87, 110, 111, 248, 251, 252, 254, 255, 427
- prosodically motivated movement (p-movement) 252
- prosody 110, 249, 256, 258, 427, 538
- Provençal 441, 476, 550, 746
- proverb(s) 283–284, 489, 606–608, 625
- proximity
 - to the addressee 134, 153, 189, 292, 385, 689
 - to the addresser 134, 153, 171, 189, 292, 385, 689
- Punic 313, 564
- purism (see also: Catalan) 488
- purpose 400–401
- purpose and result clauses 235–239
- Pyrenees 30, 54, 313, 359, 371, 372, 437, 438, 440, 441, 443, 446, 471, 518, 550, 563, 567, 599, 650, 618, 651, 670
- qualitative data 620
- qualitative exclamatives 264, 265
- quantifier(s) 131–135, 170–174, 181–190, 207–208
- quantitative data 334, 620
- quantitative distinction 137
- quantitative exclamatives 264–265
- quantitative meaning 170, 175, 188
- quantity 134, 174, 244
- Quebec 31, 616, 695, 700, 705
- question (issue)
 - language question 661
 - national question 23
 - ‘question of names’ 19, 558–559
 - *questione della lingua* 339, 491
- question (syntax)
 - echo question 263
 - tag question 611
- race 528, 726
- raising 106, 107, 449, 450, 606
- reading 32–33, 51, 86, 248, 257, 265, 274, 277, 315, 410, 463, 488, 505, 586, 638, 684, 715
- reading difficulties 86
- reading rules 87

- Real Cédula* 492, 670
 real time 405–408
 reciprocal 156, 177
 reconquest/Reconquest 19, 443–451
 reference form 142
 referent 131–132, 225–226, 276, 298, 328, 562, 569, 574
 reflexive 156, 176–177, 363
 regional varieties (see: variety, varieties)
 register (see: variety, varieties)
 relative
 – adverb 218, 221, 222, 233
 – determiner 218
 – pronoun 183, 218–222
 relative clause 218–234, 244, 277
 – free relative clause 171, 219, 226
 – in infinitive 222, 223
 – non-restrictive relative clause 219, 220, 224, 244
 – restrictive relative clause 219
 – semi-free relative clause 220, 222
 Relevance-heuristic principle 323
 Relevance Maxim 323
 religion 326, 443, 486, 509, 566, 571, 582, 726
 remoteness 134, 135
 renaissance 23, 519, 630
 Renaissance 486, 488, 489, 491, 497, 739, 740
Renaixença 6, 29, 34, 497–519, 710
 Renan, Ernest 18
 repertoire
 – functional 401
 – language/linguistic 10, 398, 403, 413, 417, 635, 671
 – lexical 4, 8, 316, 319, 331,
 – multilingual 398
 – normative 336
 – symbolic 510
 – verbal 8
 repression 31, 49, 343, 446, 546–547, 666, 711, 732
res (negative quantifier) 167, 174, 615
 resources
 – expressive 337, 606, 622
 – linguistic 423, 425, 431, 481
 – verbal 605, 612, 614, 617
 – Written Language Resources for EU Languages 429
 result construction 229, 235
 revival 5, 78, 497, 498, 503, 504, 555, 581, 589
 rhematic 237–238
 Riba, Carles 6–8, 555, 586
 Ribagorçan 153, 384, 385
 Robins, Robert H. 37, 44
 Roman empire (see: Rome)
 Romance 314, 329–338, 442, 446, 447, 448, 456, 457, 465, 478, 487, 490, 550, 565, 566, 567, 569, 570, 575, 578
 – Gallo-Romance 167, 315, 475, 490, 599
 – Ibero-Romance 167, 314, 475, 490, 599
 – Italo-Romance 599
 – language(s) 4, 25, 46, 49, 84, 88, 97, 165–167, 205, 222, 255, 258, 266–267, 279, 281, 314–342, 357, 366, 377, 381, 382, 437, 440–446, 474, 535, 599, 637, 677, 704
 – philology 19, 474
 – Romance-Arabic 570
 – terminology 331, 443
 – varieties 314
 – vernaculars 446
 – vocabulary 337
 Romanization (also: Romanisation) 21, 438–441, 562, 567
 Romantic 6, 20–22, 488, 498, 502, 507, 561
 Rome 313–314, 437–441, 473, 479, 482, 740–746
 – Holy Roman Empire 25, 485
 – pre-Roman 313, 442, 562–564
 – Roman 442, 564, 576, 577
 – Roman anthroponymy 577
 – Roman court 483
 – Roman Empire 27, 314, 440
 – Western Roman Empire 441, 485
 root
 – compounds (see: compound)
 – extension 140, 160–161, 378–379
 Rosselló 13, 471, 474, 520–521, 531, 549
 Roussillon (see also: Rosselló) 13, 438, 441–443, 487, 518, 631, 653
 routine formulae 283–284
 Rubió i Lluch, Antoni 525, 534
 Rubió i Ors, Joaquim 34, 502–504, 507
 rule (dominance)
 – Bourbon/Catalan-Aragonese/Muslim/Roman/self-rule/Visigothic rule 40, 439, 441, 566, 600, 740
 – home rule 38, 55, 76, 78, 508

- rule (grammar)
- inflectional/morphological/phonological/syntactic rule 7, 104, 109, 110, 112, 115–119, 135, 150, 161, 404, 615, 616
 - variable rule 404, 407
- Russell-Gebbet, Paul 55, 439, 441, 444, 447, 451, 463, 465, 466
- Russian 23, 30, 587, 590, 735
- Sales, Joan 555, 588, 591
- Sanchis Guarner, Manuel 43, 50, 52, 53, 55, 58, 78, 532, 555, 558
- Sant Jordi 510
- sardana 510
- Sardinia 9, 13, 316, 318, 328, 335, 373–374, 449, 472–473, 479, 527, 570, 599, 616, 619, 652, 740, 746
- Sardinian 13, 319, 335, 375, 377, 390, 476, 479, 599, 608, 652, 655–656, 676, 740
- Saroihandy, Jean-Joseph 40–43, 525, 745
- Saussure, Ferdinand de 5, 18, 44, 82
- Scandinavian 578
- Schädel, Bernhard 40–47, 745
- Schism 520
- Great Schism of Western Christianity 18, 480
- school 11, 33, 40, 407, 500, 517, 537, 540, 549, 552, 554, 556, 601, 612, 618, 638, 655, 657, 670–675, 683, 687, 724, 730, 732, 735
- schooling 33, 598, 604, 606, 638, 670
- schwa 102, 113, 144, 148, 149, 373
- scribal error 450, 451, 463
- scribal traditions 450
- segment (see also: morphological segmentation) 106, 146, 615
- self-assertion 8–9, 493
- self-hatred 603, 715
- semantic autonomy 277
- semantic basis 136, 351
- semantic content 130, 577
- semantic difference 136–137
- semantic feature 134
- semantic field(s) 568–569, 574–575
- semantic foundation 137
- semantic property 273–283
- semantic referents 562, 566
- semantics 58, 263, 313, 317, 320, 333, 352, 355, 366, 608
- Semitic 6, 599
- sentence
- boundary 123
 - copular sentence 178–181, 196–198
 - declarative sentence 122, 125, 168, 206, 214–215, 258, 302
 - exclamative or exclamatory sentence 214, 235, 247, 249–250, 263–267
 - interrogative sentence 174, 207, 214, 225, 249–250, 258–263
 - periphery 238
 - quasi-copular sentence 165, 196–198
- senyera* 510
- Sephardi 740
- Septimania 439–446
- sequence of tenses 243
- ser/ésser* (verb) 181, 187, 192, 196, 197, 198, 217, 317, 385, 407, 409, 465
- Serbia, Serbians 25–26
- Serbo-Croatian 25–26
- sermo rusticus* 491
- sermo urbanus* 491
- seu, seva/seua, seus, seves/seues* (third person stressed possessive) 155, 172, 389, 460
- sex 137, 321, 409, 609, 640
- shared knowledge 288–290, 308
- Sharq al-Andalus (see also: Kingdom of Valencia) 472–473
- sí* (adverb) 188, 206–207, 242, 299
- sibilant 104, 106, 107, 108, 110, 114, 117, 119, 122, 151, 161, 376, 378, 383, 385–391, 453
- Sicilian 476, 479, 599
- Sicily 318, 472–473, 479, 740
- simple lexical base 574
- simple (verb) form 142, 158, 201, 292
- simplification 107, 110, 385, 448, 453, 454, 622
- simultaneity 140, 190, 199, 228
- Sinclair, John 421
- singular 136–138, 139–140, 136–163
- Slavic 578
- social change 713, 725
- social class 336, 339–341, 399, 409, 630, 633, 640, 713
- middle class 25, 29, 30, 617, 669, 711, 713
- social cohesion 724, 727–729, 731, 734, 735
- social networks 337, 598, 618, 619
- social practices (see: practices)
- social status 324, 342, 440, 714
- Societat Catalana de Llengua i Literatura (SCLL) 747

- Societat Catalana de Terminologia (SCA-TERM) 701
- Societat d'Onomàstica 56, 562
- sociolect 321, 342, 407
- sociolinguistic 2, 9–10, 19, 53, 60, 287, 312–315, 326, 331, 335, 336, 340, 371, 391, 397, 398, 401–406, 412, 423, 440, 479, 482, 534, 597–598, 605, 616, 631–633, 635, 639, 641–642, 651, 695, 700, 709, 710, 714, 725, 730
- sociolinguistic variables 404
- sociolinguistics 18, 402–403, 405, 410, 695, 696, 725
- diachronic sociolinguistics 311, 337
 - variational sociolinguistics 340, 403
- sociopragmatic variables
- age 342, 404–409, 427, 427, 609, 640
 - ethnic group 9, 27, 399, 404
 - gender 149, 306, 404, 409, 640, 713
 - socioeconomic class 404, 409, 640, 646
 - style 340, 404
- Solà, Joan 49, 53–61, 529, 556–557, 710
- Soldevila, Ferran 445
- son, sa, sos, ses* (third person unstressed possessive) 460, 171
- Song of Songs* 6, 586
- Spain 10, 12, 32, 37, 47–48, 55, 59, 332, 371–372, 401, 445, 487, 498–500, 505–507, 539, 553–554, 600, 602–603, 614, 616, 618, 631, 633–636, 643, 649–651, 658, 666, 670, 677, 699, 704, 711, 724, 726, 729, 730, 741
- Spaniards 20, 28, 500, 740
- Spanish
- Academy (*Real Academia Española*) 492, 499, 538
 - Autonomic State 651
 - Civil War 6, 34, 47, 49, 51, 77, 343, 545, 546, 586, 618, 634, 676, 680, 742, 745
 - conservative press 671
 - Constitution 8, 55, 498, 554, 651
 - Cortes 633
 - dictatorship(s) 5, 49, 77, 287, 316, 343, 538, 546, 547, 548, 550, 551, 554, 581, 586, 588, 635, 643, 670, 678, 698, 711, 746
 - First Spanish Republic 498
 - government 729
 - hegemony 581, 587, 712, 715
 - literature 32
 - market 33
 - military 499, 506, 508, 525, 538, 546, 615, 633, 740
 - Parliament 540
 - periodicals 504
 - political parties 712
 - Second Spanish Republic 49, 77
 - State 8, 415, 498–499, 506, 549, 551–552, 554, 632, 710, 716–717, 731, 734
 - transition to democracy 60, 651, 672, 710
- Spanish (language) 523, 725, 728, 730
- acquisition 501, 714
 - as a foreign language 407, 703
 - classes (also: courses) 730–731
 - colloquial expression 625
 - command of Spanish 413
 - idiom 625
 - interference on Catalan 7, 59, 82, 227, 342, 344, 407–408, 413, 415, 490, 622, 624, 701
 - language of education 499
 - learning 633, 728–729
 - lexical-semantic changes 344
 - lingua franca 728, 735
 - literary language 25, 32–33
 - national language (see also: Castilian) 499, 511
 - naturalization of Spanish 344, 654
 - official language (see also: Castilian) 536, 550, 600, 665, 667
 - Peruvian Spanish 256
 - regional Spanish 500
 - Spanish neology network 704
 - speakers 727
 - standard Spanish 500
 - translations 503
 - variety 44, 256
 - word-formation patterns 351
 - written language 31, 618
- speaker
- elderly speakers 326, 388, 409, 732
 - fluent speaker 622
 - L1 speaker 643, 645, 671, 679
 - L2 speaker 624, 712
 - new speaker 10–11, 636, 644–645, 728
 - (non-)native speaker 10, 46, 86, 352, 597, 603, 618, 636, 639–646, 674, 712, 728
 - number of speakers 82, 423, 599
 - semispeaker 622–624
 - speaker's attitude 133, 324

- speaker's point of view/perspective 205, 297
- speaking 20, 22
- speech
 - indirect speech 608
 - reported speech 201, 609
 - speech and voice disorders 47–50
 - speech event 410, 602, 609
- spelling (see also: orthography) 6–8, 11, 38, 43, 52, 81–90, 93–97
 - rules 11, 358, 522, 528, 531, 533
- standard language (see: language)
- standardisation (also: standardization) 3–9, 38, 57, 62, 63, 107, 336, 342–343, 382, 398, 411–415, 421, 491, 500–503, 511, 532–540, 548, 601, 661, 684, 694
- State 24–28, 487–488, 492, 498, 618, 650, 662–666, 670, 716, 724, 748
- status planning (see also: language planning) 9, 545, 630
- Statute of Autonomy
 - of Aragon 659
 - of Catalonia 540, 545, 554, 651, 663, 664, 671
 - of Murcia 658
 - of the Balearic Islands 651, 662
 - of Valencia 559, 651, 660
- Statute of the Municipality of l'Alguer 656
- statute(s) 651, 656, 660, 662, 664, 746
- stem 105, 109, 115, 119, 120, 122, 130, 142, 145, 147, 158–159, 464
- stigmatisation (also: stigmatization) 492, 619
- stress
 - assignment 121–122
 - oxytone(s) 113, 118–121, 150
 - paroxytones 118–121, 389, 456
 - phrase stress 108, 110
 - preaccented suffixes 113, 121,
 - proparoxytones 119–122, 378, 387–388, 456
- style 32, 243, 480, 481, 530, 531, 532, 548, 586, 684, 686
- stylebooks 557, 684–691
- stylistic criteria 686, 691
- subclass 131–133, 135–140, 160, 261
- subject 122–123, 131–132, 137–138, 141, 155, 165–166, 175–176, 183, 188, 194–196, 214–218, 220, 223–224, 236, 239, 250–254, 255, 259, 262, 279–280, 296–298, 353, 356–358, 363
 - animate subject 197
 - covert subject 165, 184
 - dislocated subject 251, 255, 265
 - elliptical subject of the infinitive clause 215–216
 - null subject 231
 - of the sentence 168, 192, 194, 195, 205
 - omission of the subject 166
 - overt subject 165
 - position 175, 247, 291, 248, 291, 298
 - post-verbal position 166, 178, 194, 216, 248
 - preverbal 266
 - subject-verb inversion 266–267
- subjective person 143
- subjectivity 223, 288, 298, 732
- subordinate clause
 - infinitive subordinate clause 215–216, 239
 - subordinate content clause 180, 214, 223–224, 227
 - temporal subordinate clause 227–229
- subjunctive
 - in adjunct constructions 226
 - in content clauses 214–218, 223–224, 227
 - in independent sentences 224
 - in relative clauses 167, 218
- substrate 440–442, 474, 563, 569
- suffix 84, 351, 575, 578
- suffixation 119, 353, 356–359, 364, 568–575
- superlative construction 229, 234
- superstrate 440, 449
- suprasegmental features 57, 121–126
- surname(s) (see: names)
- syllable 22, 87–92, 105–125, 147–149, 157, 162–163, 367, 373–390, 449–450, 452–456
 - coda 107, 108, 114–117
 - margins 106, 108, 113
 - nucleus, nuclei 106–107, 110
 - onset 105–111, 116
- syllabification 103, 107, 111, 150, 161
- symbolic 26, 84, 96–97, 416–417, 478, 499, 508–510, 518, 556, 680, 698, 730
- syncretic, syncretism 129, 144, 162–163, 255, 381–382
- syntactic
 - autonomy 129
 - component 137–138
 - constituent 133, 168, 183–184, 187–192, 206–207, 211–212, 219, 230–231, 250–254, 258, 274, 300, 331

- feature 131
- function 133, 138, 177, 214, 219, 222
- (in)flexibility 273–278
- property 187, 283
- relation(ship) 137
- role 136
- structure 213, 230, 251, 457, 685
- syntactic-phonetic factors 456

- tag (see: question, codeswitching)
- tagging 422, 428–429
- tail 256–257
- temporal construction 227–229
- tenor 400–401
- tense 136–143, 146, 196, 198–202, 465, 624
 - compound 137–138, 158, 199, 201, 217, 475,
 - conditional 144, 158, 159, 198–203, 243, 381, 465
 - conditional perfect 198, 199–203, 243
 - future 121, 139, 140–143, 158–159, 198–199, 201, 203, 205, 242, 243, 465,
 - future perfect 199–201
 - imperfect 144, 158–160, 198–199, 202–203, 243, 381, 389, 391, 464–465
 - indicative tenses 141–144, 158, 160–161, 625
 - past 140–146, 160, 166, 202,
 - past anterior 199–203
 - periphrastic past 158, 166, 198–199, 203
 - pluperfect 198–200, 203, 243, 466
 - present 143, 161, 198, 202, 366, 373, 379, 380–386, 389–391, 625
 - present perfect 158, 199–200, 203, 466
 - simple past 158, 160, 166, 199, 203,
 - subjunctive tenses 144, 158, 160, 162, 203, 389
- term 6, 7, 46, 317, 324, 326, 332, 339, 443, 572, 608, 631, 683, 690, 695–700, 703
- TERMCAT 2–3, 58, 319, 344, 686, 699–704
- terminological normalisation 696, 697, 700–701
- terminology
 - applied terminology or terminography 696–698
 - Catalan terminology (see: Catalan)
 - institutionalisation of terminology 693, 694, 699
 - management 696
 - normalised terminology 698, 700, 703
- territorial expansion 19, 437, 474, 567, 573, 577
- teu, teva/teua, teus, teves/teues* (second person singular stressed possessive) 154, 155, 171, 460–461
- text
 - as a building block of shared culture 608
 - text types 400–401, 426
- Textsprache* 599
- thematic 237–238, 256, 277, 280
- theme (grammar) 194–195
- ton, ta, tos, tes* (second person singular unstressed possessive) 154
- toponym (see also: Catalan) 452, 456, 459
- toponymic studies 54
- transitive (see: verb)
- translanguaging 725
- translation 581–590, 696, 703–705
 - and literary language 5–8
 - contemporary translation 588–589
 - false friends 625
 - from Catalan 590–591, 748
 - into Catalan 297, 319, 331, 402, 476–478, 582–583, 584–588, 589, 655
 - literal translation 294–298
 - machine-translation 624–625, 744
 - medieval translation 582–583
 - modern translation 583–584
 - strategy, strategies 293, 296
- transparency 357, 678
- transposition 352–358
- Trastámaras 473–479
- treaty of Corbeil 474
- treaty of the Pyrenees 315, 372, 487, 583, 600
- truncation 368
- tu* (second person singular strong pronoun) 139, 155, 175, 291, 321, 461
- type(s) of discourse 422, 427

- Überdachung* ('roofing') 501
- Ullastra, Josep 38, 491
- Umbrian 440
- un, una, uns, unes* (indefinite article) 153, 170
- Unamuno, Miguel de 21
- unit
 - accentual unit 155
 - communicative unit 288
 - lexical unit 55, 282, 695
 - linguistic unit 288, 294, 612

- phonological unit 133
- stress unit 154
- translation unit 430
- United Nations 34
- unity (of the language) 8, 9, 343, 415, 492, 493, 531, 662
- Universal Declaration of Linguistic Rights* 716
- unstressed (see: possessive; pronoun; vowel)
- urban nomenclature 574
- Urnenfelder* 313, 563
- us, vos* (second person plural weak pronoun) 133, 139, 156–157, 176, 390, 461–462
- utterance 133, 134, 229, 248–252, 283, 291, 300–301, 307, 605, 606, 611, 613

- va* (past auxiliary verb) 146, 158, 159, 161, 166
- Valencia
 - *Generalitat Valenciana* 22
 - Kingdom of Valencia 19, 28, 31, 315, 335, 472–476, 486, 493, 524, 630
- Valencian
 - Central Valencian (*Apitxat*) 87, 104, 113, 383–385, 451, 506
 - codification and standardisation 62, 661
 - idiomatic particularity 532
 - idiomatic secessionism 558
 - official status 559, 660, 661
 - standard language 661
 - *valenciana prosa* ('Valencian prose style') 340, 481
 - variety 25, 61, 62, 148, 151, 385, 417, 715
- Valencian Community 56, 563, 566, 651, 658
- Valencian Country 9, 13, 19, 24–25, 33, 52–53, 300, 372–374, 382, 389, 410, 412–414, 617, 634–646, 671, 677–678, 685
- Valera, Éamon de 18
- Vallverdú, Francesc 57–60, 63, 588
- variation
 - and change 312–321
 - diachronic 319
 - dialectal 87, 124, 161, 196, 253–255, 258, 266, 332–336, 339–340, 391, 399–400, 526
 - diaphasic 337, 339, 340
 - diastatic 340, 440, 557
 - diatopic 239, 440, 476–477, 490, 512, 530, 557
 - functional 328, 336–337, 397–402, 411, 416
 - geolinguistic 56, 62, 378
 - interlinguistic variation 398
 - intralinguistic variation 398, 401
 - lexical-semantic variation 332–334, 341
 - regional 57, 325, 530
 - social 336, 340
 - stylistic 340, 400, 404–405, 409–415
 - through time and space 312
- variationism 397, 403–404, 408, 410–411
- variety, varieties
 - cultured 491
 - dialectal 29, 145, 288, 315, 321, 328, 334, 336, 399, 522
 - functional 321, 398, 399–401
 - geolectal, geographical 61, 321, 477
 - prescriptive 87
 - regional 325, 687
 - register 184, 221, 288, 315, 317, 321, 327, 328, 337–342, 368, 399, 400–411, 477–481, 500, 529, 557, 684–690, 714, 732
 - social 321, 399
 - sociolect 321
 - standard 147, 398–401, 409, 687
- Veny, Joan 53, 56, 62, 312, 332, 376, 378, 379, 383, 384, 391
- verb
 - auxiliary verb 158, 166, 191, 204–205, 407, 465
 - causative verb 195
 - copular verb 132, 179, 192, 196–198, 218
 - ergative verb 194
 - first conjugation verb (see also: conjugation) 378
 - inchoative verb 140, 623
 - intransitive verb 192–195, 363, 465
 - irregular verb 44, 146, 158
 - light verb 192, 272, 327
 - linking verb 131
 - measure verb 193
 - meteorological verb 192
 - non-copular verb 131
 - perception verb 193, 217–218, 306
 - predicative verb 132
 - quasi-copular verb 179, 192, 196–198
 - regular verb 44, 146, 158
 - second conjugation verb (see also: conjugation) 140, 159, 161, 463, 624
 - third conjugation verb (see also: conjugation) 140, 159, 160–162, 378, 380–381, 463, 624

- third inchoative conjugation verb (see also: conjugation) 140, 159, 160
- third pure conjugation verb (see also: conjugation) 140
- transitive verb 131, 178, 192–195, 217, 363, 465
- unaccusative verb 178, 192–195, 248
- unergative verb (see also: intransitive verb) 193
- verb phrase(s) (VP) 168, 191, 206, 385
- verbal conjugation (see also: conjugation) 142
- verb(al) form 143, 166, 186, 625
- verbal function 624
- verbal inflexion 50
- verbal mood (see also: mood) 141
- verbal periphrasis/periphrases 192, 204–206, 611
- verbal person (see also: person) 146
- verb(al) tense (see also: tense) 137–138
- verbalisation 356–358
- Verdaguer, Jacint 7, 8, 503–504, 530, 534, 535, 548, 590
- Viciano, Rafael Martí de 38
- Vilanova, Arnau de 29, 478, 480
- Vilar, Pierre 488, 601, 617
- Visigothic 441–449, 565, 576–577
- vocabulary 4, 32, 39, 53, 311–344, 475, 481, 503, 530, 622, 685, 724, 740
- vocalization
 - of syllable-final and word-final /ð/ to -u (-[w]) 452
 - of syllable-final /l/ 453
- voice
 - disorders 47
 - recognition 425
 - recording 378, 425–427, 607–614
- voice(s) 40, 610–613, 712, 728
- voiced 87–89, 102–104, 108–109, 113–119, 378, 385, 387, 391, 405–406, 451–454
- voiceless 48, 88, 95, 120, 378, 389, 406, 407, 453
- voler* (irregular verb) 146, 206
- Volk* 506, 509
- Volkgeist* 21
- vosaltres* (second person plural strong pronoun) 133, 139, 155–156, 175–176, 291
- vostre, vostra, vostres* (second person plural stressed possessive) 155, 172, 291, 460
- vowels
 - atonic (see also: unstressed vowel) 442, 450, 462, 474–475, 491
 - close vowel 38, 91, 106–113, 474–475
 - close-mid vowel 91, 102, 121, 615
 - elision 93, 110–113
 - epenthesis 107, 109, 110–114, 119–120
 - front vowel(s) 88, 102, 103, 388
 - harmony 114–116, 375
 - open vowel 93, 121, 449
 - open-mid vowel 91, 102, 115, 615
 - post-tonic 375
 - pretonic 116, 376
 - reduction 7, 38, 84–87, 91, 332, 389, 391, 448, 615–616
 - stressed vowel(s) 91, 92, 101–102, 111–116, 121–122, 144, 146, 159, 161, 374–377, 380–381, 386–387, 390, 442, 447–450
 - stressed /ə/ 102, 376, 377, 378
 - support vowel 148–153, 157, 161, 380
 - syncope 111, 442, 451, 454
 - thematic 159, 160, 161, 381, 624
 - unstressed 38, 87, 102, 110, 115, 148, 164, 328, 374, 375, 380, 388, 391, 442, 449
 - vowel system 102, 115, 388
 - weakening 103, 117, 118, 449
- War of the Spanish Succession 315, 316, 511, 740
- ways of speaking 410, 613, 714
- Weinreich, Uriel 403, 598, 605
- wh-exclamatives 265
- wh-interrogatives 253, 259, 261, 263, 265
- word
 - classes 130–136, 354, 359, 362
 - cognate forms 194, 322, 327, 353
 - complex word 272, 281, 352
 - compound word 85, 110, 115, 122, 130, 148, 187, 272
 - constituent 129–130
 - derived word (see also: derive) 103, 115, 148, 151, 272, 374
 - inflected word 130, 147, 159
 - lexical word 142, 362
 - simple word 130, 272
 - stress (see also: word-accent) 106, 122, 154
 - wh-word 258, 259, 261, 263, 265
 - word-accent (see also: word stress) 155
 - word-boundary 108

- word-final position 103–117, 135, 155, 161–162, 381, 390
- word-formation
 - deadjectival 187, 353–362
 - denominal 272, 356, 364
 - deverbal 353–358, 363–364
- word order 168, 247–254, 258, 263–267
- World War II 48, 535, 547, 549, 617, 745
- writing 84, 85, 104
 - alphabetic writing systems 83, 93
 - conventions 84–85, 87, 90
 - difficulties 86, 97
 - duality of writing systems 86
 - practices (see: practices)
- written form(s) 84, 104, 319, 506, 601, 630
- written language(s) 28, 31, 33, 84, 158, 416
- Wüster, Eugen 693–695
- Xipella* 383
- yes/no interrogatives 258–260
- Yiddish 20, 23
- yod 146, 381, 391, 406, 409, 410, 463
- zero-derivation 353, 354
- zero morph 354