

Romance Languages and Linguistic Theory 2018

EDITED BY
Sergio Baauw
Frank Drijkoningen
Luisa Meroni

JOHN BENJAMINS PUBLISHING COMPANY

CURRENT ISSUES IN LINGUISTIC THEORY

AMSTERDAM STUDIES IN THE THEORY AND HISTORY
OF LINGUISTIC SCIENCE – Series IV

ISSN 0304-0763

General Editor

JOSEPH C. SALMONS

University of Wisconsin–Madison

jsalmons@wisc.edu

Founder & General Editor (1975-2015)

E.F.K. KOERNER

Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin

Current Issues in Linguistic Theory (CILT) is a theory-oriented series which welcomes contributions from scholars who have significant proposals that advance our understanding of language, its structure, its function and especially its historical development. CILT offers an outlet for meaningful contributions to current linguistic debate.

A complete list of titles in this series can be found on benjamins.com/catalog/cilt

Editorial Board

Claire Bowern (New Haven, Ct.)

Alexandra D'Arcy (Victoria, B.C.)

Sheila Embleton (Toronto)

Elly van Gelderen (Tempe, Ariz.)

Iván Igartua (Vitoria-Gasteiz)

John E. Joseph (Edinburgh)

Matthew Juge (San Marcos, Tex.)

Danny Law (Austin, Tex.)

Martin Maiden (Oxford)

Martha Ratliff (Detroit, Mich.)

Klaas Willems (Ghent)

Volume 357

Sergio Baauw, Frank Drijkoningen and Luisa Meroni (eds.)

Romance Languages and Linguistic Theory 2018
Selected papers from 'Going Romance' 32, Utrecht

ROMANCE LANGUAGES AND LINGUISTIC THEORY 2018

SELECTED PAPERS FROM
'GOING ROMANCE' 32, UTRECHT

Edited by

SERGIO BAAUW
FRANK DRIJKONINGEN
LUISA MERONI

Utrecht University

JOHN BENJAMINS PUBLISHING COMPANY
AMSTERDAM & PHILADELPHIA



The paper used in this publication meets the minimum requirements of the American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

DOI 10.1075/cilt.357

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2021043995 (PRINT) / 2021043996 (E-BOOK)

ISBN 978 90 272 1012 8 (HB)

ISBN 978 90 272 5829 8 (E-BOOK)

© 2021 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Company · <https://benjamins.com>

Table of contents

Introduction	1
<i>Frank Drijkoningen, Sergio Baauw and Luisa Meroni</i>	
CHAPTER 1	
Processing clitic pronouns outside coargumenthood	11
<i>Valentina Brunetto</i>	
CHAPTER 2	
Infinitival complement clauses: Data from L2 acquisition of European Portuguese	25
<i>Aida Cardoso, Inês Duarte and Ana Lúcia Santos</i>	
CHAPTER 3	
Focus fronting vs. <i>wh</i> -movement: Evidence from Sardinian	49
<i>Silvio Cruschina</i>	
CHAPTER 4	
The varieties of temporal anaphora and temporal coincidence	71
<i>Hamida Demirdache</i>	
CHAPTER 5	
The structure and interpretation of ‘non-matching’ split interrogatives in Spanish	97
<i>Olga Fernández-Soriano</i>	
CHAPTER 6	
Differential object marking and scale reversals	117
<i>Monica Alexandrina Irimia and Anna Pineda</i>	
CHAPTER 7	
Contact phenomena: The I-language of a bilingual	131
<i>Luis López</i>	

CHAPTER 8	
- <i>ŋ</i> plurals in North Lombard varieties: Differential plural marking and phases	151
<i>M. Rita Manzini, Leonardo M. Savoia and Benedetta Baldi</i>	
CHAPTER 9	
Brazilian and European Portuguese and Holmberg's 2005 typology of null subject languages	171
<i>Ana Maria Martins and Jairo Nunes</i>	
CHAPTER 10	
Aspect in the acquisition of the Spanish locative paradigm by Italian L2 learners	191
<i>Silvia Perpiñán and Rafael Marín</i>	
CHAPTER 11	
Catalan nativization patterns in the light of weighted scalar constraints	205
<i>Clàudia Pons-Moll and Francesc Torres-Tamarit</i>	
CHAPTER 12	
Temporal marking and (in)accessibility in Capeverdean	225
<i>Fernanda Pratas</i>	
CHAPTER 13	
Very extracted: On old Italian <i>molto</i>	249
<i>Silvia Rossi and Cecilia Poletto</i>	
CHAPTER 14	
On adverbial perfect participial clauses in Portuguese varieties and British English	263
<i>Purificação Silvano, António Leal and João Cordeiro</i>	
CHAPTER 15	
<i>Craindre</i> ("fear") and expletive negation in diachrony	287
<i>Chloé Tahar</i>	
CHAPTER 16	
Fission in Romance demonstrative-reinforcer constructions	303
<i>Silvia Terenghi</i>	
Index	317

Introduction

Frank Drijkoningen, Sergio Baauw and Luisa Meroni
Utrecht University

For venue XXXII, the community of Going Romance returned to its historical base, the city of Utrecht, at the research institute now known as UiL – OTS. The general main session, as well as two special sessions took place in December 2018. The special sessions were related to two large research projects granted to Romance specialists at UiL-OTS: one on “Romance varieties in (micro-)contact and linguistic theory” and another entitled “GREAT 2018: Going Romance (Re-)Explores Aspect and Tense”.

From its inception in the 1980s, the circuit of Going Romance adheres to the generativist point of view that language in general as well as specific languages and dialects are best studied by postulating an overarching distinction between universal general (invariant) properties of languages and properties that are liable to differences and variation in space and time. Beyond the obvious fact that each language has its own lexicon or vocabulary, a dominant idea was to formulate parameters in syntax in order to account for differences among languages. Some standard parameters include the pro-drop parameter, the OV/VO parameter, or the V2 parameter. One of the contributions to linguistic theory was that parameters could cut across traditional language family classifications often based on historical developments of sounds and vocabularies. For example, Modern French is not a pro-drop language although French is part of the Romance language family.

One of the other important contributions to linguistic theory was that the very idea of having parameters could provide a solid basis for the quality (speed and level of proficiency) of the acquisition of syntax by children: if the invariant property that verbs combine with nouns is innate (or an early process) what needs to be learned is reduced to a binary choice (noun after verb or verb after noun) that will be learnable for the child on the basis of some simple primary data in its environment (two words phase). Within this approach it is natural to proceed by studying actual first language acquisition and also other language acquisition processes in order to determine for example how second language learners would learn the non-invariant (different) properties of the non-native language and un-learn

the parameter settings (loosing previously set properties). As there was a general idea that these parameters were sound and relevant for acquisition processes, they paved the road for much more research paying attention to actual concrete first language acquisition processes as well as specific issues in early and late second language acquisition (or learning).

This approach has been successful in the sense that it provided basic descriptions of morpho-syntactic differences between languages and their function for acquisition processes. Nevertheless, since the 1980s there have been at least three different developments that are useful to mention here. It is not the case that the very idea of splitting invariant properties from variable language properties is at stake, rather that parameters in this particular conception do not appear to be enough to account for the variation that is actually attested.

The implementations of the invariant properties had a bias towards Indo-European languages, perhaps due to the dominant position of Western societies in linguistic theorizing and/or the importance of English as a language of universal trade and communication. This led to the idea of macro-parameters, higher level differences, larger than just inside one language family (e.g., ergative and polysynthetic languages). And then the idea of micro-parameters is not far away either, parameters to just state the tiny differences among different dialects of one language (Italy as the core example of a country in which it is a general theme); in particular, without a normative attitude (that sometimes seems to prevail in France and Spain), such variation is found almost anywhere. This development does not conflict with the parameter concept itself, but rather enlarges its scope. The discussion of micro-parameters in this sense has led to a renewed attention to language contact and situations of natural bilingualism, e.g., native speakers of a dialect or a minority language in situations (like schools or administrations) that make use of the language of the political nation.

The second development is a somewhat more intrinsic inconvenient property of parameters, linked to the specific way of using parameters for describing differences at the level of languages. For example, although one could very well speak of a V2 parameter in order to account for a relevant syntactic property of word order in German or Dutch as opposed to English or French, it is inconvenient that both English and French contain a subset of constructions that exhibit the relevant V2 property (respectively Subject-Aux inversion and Subject Clitic inversion). This seems to entail that the V2 property (independently of its exact formal characterization) is a property of particular syntactic constructions and therefore not only of entire languages. Likewise, the parameter for polysynthetic languages (a macro-parameter) is useful for the explanation of English verbal compounds. Similarly, the position of the subject in Locative Inversion in English seems to share important properties with the postverbal subject in pro-drop languages. Or, still

another example, although Dutch is not a pro-drop language, there are some constructions allowing for exactly this (expletive pro drop). All these examples seem to show that there exist syntactic properties that are termed parametric that can be used to construct sentences but that the properties themselves are not literally defining language differences only; they are also defining construction differences. This does not affect the idea of parameters as a way of stating differences or variation, but it does affect the idea of parameters as the sufficient means to define a language in opposition to other languages. It might be the case that a language can be characterized by a full set of construction-bound set parametric properties, changing the status of parameters.

A third development has been central at the abstract level, the relevance of morphology in syntax. This had been noted but its importance probably had been underestimated. It is generally clear that languages that allow pro drop are languages with rich markings on the verb, that is: morpho-syntactic properties intervene. For the macro-parameter concerning polysynthesis, the exact same idea holds: it interacts heavily with morphology: polysynthesis essentially is making morphologically complex words in situations where Indo-European languages make syntactically complex combinations of simple words. Dialectal variation is often micro-level variation in the field of indications for phi-features and other inflection or agreement related properties, that is, once again, morphological variation. This development has led to the idea that differences in morphology are considered to be the prime locus for the determination of syntactic variation. As said, it is rather obvious that languages differ in their lexicon; what is more or less new is that there are good reasons to believe that languages differ in their morpho-syntactic properties and that this might well be sufficient for variation in morpho-syntax. Since the start of the minimalist program this led to issues of strength/weakness of possibly abstract features ('abstract' meaning 'possibly without an actual morpheme'), including the possibility that features guide the actual movement steps (criterial usage of features). In general, it seems that nowadays morphological analysis is intrinsically intertwined with syntactic structure building, where the actual morpho-syntactic forms may be dealt with relatively late (some spell-out point and/or Distributed Morphology) or relatively early (minimalist lexicalist approaches with checking procedures and probing). There are relations between words, morphemes and features, but these appear not to be one-to-one (fusion, fission, polysynthesis, phonological emptiness, semantic emptiness (minus interpretable as non-interpretable)).

A good number of the papers related to syntax and morphosyntax in this volume can be characterized in terms of these ideas and developments. Papers in other linguistic subdisciplines can be placed along the same lines. In phonology, Optimality Theory also deals with variation, as different rankings of invariant simple constraints. Assuming that logical reasoning is universal and invariant, theories

of formal semantics also take into account language variation. Simple examples involve differences in verb morphology, such as the existence of the progressive in English or the existence of subjunctives in Italian; a more complex example would be that the English present perfect is not always semantically identical to the corresponding French forms. With respect to the general idea of the minimalist program that morphology is crucial in understanding variation, this thus appears to hold also for the semantic interpretation of the morpho-syntactic forms (linguistic relativity beyond the lexicon). In that sense, a universal feature [+interpretable] does not entail an actual interpretation. Finally, there is the field of information structure; languages may and do vary in the positions that are available in syntax for the transmission of new or old information, but it is also argued that information can be transmitted without making use of syntactic mechanisms, and that phonological issues may intervene too (one might (today still) differ of opinion on whether these notions are relevant for syntax proper or not).

In what follows we indicate ways in which the papers selected for this volume contribute to these different research issues. The reader might note overlap (copying actions) with respect to the abstracts that precede the papers in the volume.

Based on data from Brazilian and European Portuguese, Ana Maria Martins and Jairo Nunes (Chapter 9) rethink Holmberg's characterization of partial vs. consistent null subject languages (NSL). The paper proposes that, rather than overt morphological distinctions, what is relevant for null subject licensing is the underlying feature specification of the verbal inflection, after agreement between T and a pronominal subject (relevance of person/number/gender/Case features). They argue that thus only close inspection of the pronominal and agreement systems of individual NSLs permit an adequate characterization of them, for the same language may behave as a 'partial', 'consistent', or 'radical' NSL depending on the morphological feature specification of its nominative pronouns and T heads. This paper also illustrates the progress and refinements that are made in various ways with respect to the simple 1980s parameter [yes/no pro-drop].

Based on data from Sardinian, Silvio Cruschina (Chapter 3) challenges the commonly held view that focus fronting exhibits similar properties to *wh*-movement. The syntactic parallelism between the two types of movement has been supported by the semantic analyses of *wh*-questions that assume that when *wh*-words function as interrogative operators, they are inherently focal, by itself a highly attractive picture of the relationship among *wh*-words, focus, and movement. His paper claims that *wh*-phrases are not inherently focal. The results of a prosodic production experiment on the distribution of the nuclear pitch accent in Sardinian *wh*-questions, together with the syntactic properties related to the asymmetry between direct and indirect *wh*-questions, form the empirical basis for this claim. This paper can also

be considered illustrative for the (renewed) interest in information structure that integrates properties of experimental phonological nature.

In investigating spoken Peninsular Spanish, Olga Fernández-Soriano (Chapter 5) isolated a particular way of phrasing a question, exhibiting what she calls ‘Non-Matching’ Split Interrogatives. ‘Split interrogatives’ (SI) are *wh*-questions followed by a phrase that constitutes a possible answer, the ‘tag’. The overall structure is interpreted as a yes/no question (as in *what did John bring, a book?*). In standard cases, the tag matches the (case and thematic) features of the *wh*-element. Special then are ‘Non-matching Split Interrogatives’ (NMSI); in these cases, the *wh*-element and the XP in the tag may not match; instead, it is the dummy (neuter) *qué* ‘what’ that heads the *wh*-clause. She proposes a biclausal analysis involving an ellipsis process. To support this hypothesis, she argues that there is a form-meaning mismatch. This paper also testifies to the fact that despite lots of descriptive and theoretical progress in linguistic theory, there still are new data that are not yet fully accounted for, even in the field of questions and *wh*-elements.

Differences in actual morphological marking are central to the paper by Rita Manzini, Leonardo Savoia and Benedetta Baldi (Chapter 8). They provide a morphological analysis of the North Lombard feminine plurals and focus on variation with *-η* plurals. They argue that, at the syntactic level, the relevant varieties are characterized by the phenomenon of Differential Plural Marking, whereby phasal domains have different realizations of plural morphology on the head of the phase and on the complement of the phase. They provide an account of DPM based on the assumption that under the PIC the complement of the phase and its head are externalized separately. They show that this has consequences for clitics as phasal heads as well as for object agreement with participles and with finite verbs. This paper provides a prime example of how seemingly low-level dialectal variation can bear on the technical properties of the procedures of syntax (phases).

The paper by Silvia Rossi and Cecilia Poletto (Chapter 13) also shows that seemingly low-level variation can bear on syntax, but this time via the path of historical variation. They consider cases of extraction of the degree word *molto* ‘very, much’ from its modifier position within an AdjP in Old Italian. Such cases are reminiscent of Left Branch Extractions, but, differently from what happens with adverb extraction in Slavic, the conditions under which *molto*-extraction is possible are very restricted: *molto* can be extracted only when the AdjP is (or modifies a nominal expression) in post-copular predicative position but not outside a fully-fledged DP. They propose that the reason that the structure is so restricted has to do with the presence/absence of a phase boundary, while the reason that this has gone lost in modern Italian is the loss of the target position in the CP layer. This allows them to show that the loss of the Verb Second property (V2) in Italian has fine grained consequences in unexpected domains like quantifier extraction.

In her paper Silvia Terenghi (Chapter 16) proposes a new approach to Romance demonstrative-reinforcer constructions. Her account is based on a binary valued feature system for deictic person and can also illustrate the importance of features (rather than lexical items *per se*). Looking at data from Romance varieties, some (implicit) shortcomings of previous accounts can be repaired via a morphological operation: Fission. Via Fission a featural reason is given to ensure deictic compatibility, and a new approach to the category of reinforcers is put forward. This research is embedded in one of the widely used modeling devices of morphology, Distributed Morphology.

The paper by Monica Alexandrina Irimia and Anna Pineda (Chapter 6) focuses on some problematic aspects related to the diachrony of differential object marking in Old Catalan and Old Romanian (11th to 17th centuries). Corpus data from both languages reveal two unexpected facts: First, there is a prominence of the 3rd person, to the exclusion of 1st and 2nd person, contrary to what the Animacy/Person scale would predict; Second, differential marking appears to be present on nominals and especially proper names, to the exclusion of pronouns, contrary to the Specificity/Definiteness Scale. The account that they propose for these types of scale reversals builds on the idea that languages can have more than one differential object marking strategy, as well as more than one type of structure for pronouns and animate nominals. This paper also shows that the synchronic conception of parameters and their link to actual morphological marking can also be used to account for variation in time (although without thereby explaining diachronic developments).

Variation in time is also the topic of the paper by Chloé Tahar (Chapter 15) on *craindre* (“fear”) and expletive negation. She argues that in Early Old French, *craindre* can embed negative imperatives introduced by *ne*. As the use of the complementizer *que* later generalizes, *craindre* can no longer embed speech acts. She proposes that, as a result of this change in the complementation system, *ne* became a complex complementizer together with the complementizer *que* (*que ... ne*). She argues that, if *ne* changes syntactic status, becoming the fossilization of prohibitive negation, it retains its speech-act semantics in the earliest stages of the history of French, and thus constrains the interpretation of *craindre*. Building on an independent proposal that ‘fear’ verbs are hybrid attitude verbs, featuring both a doxastic and a (dis)preferential component, this paper argues that these two components are conveyed by different layers of meaning. In the earliest stages of French, expletive negation placed semantic restrictions on the main verb, which are met when the interpretation of *craindre* foregrounds the (dis)preferential component.

Thus, the mere notion of variation in the actual interpretation of a negative element reveals the intrinsic need for independent semantic analysis. In a sense, French shows that you can negate without actually negating; some morpho-syntactic negative elements can be void of negative meaning. This characterization seems to

mean that a property of the type [+interpretable] that a syntactician would generally associate to Neg (or NegP) is not as straightforwardly interpretable by semanticists as one might hypothesize. In general, the interpretability feature itself has a good function for TENSE being different from AGR in syntax and syntactic theory, but temporal interpretation actually faces the same challenge: semantically interpretable yes, but what is the actual semantic interpretation?

Historically, the actual temporal interpretation of TENSE and similar morpho-syntactic elements (PERF, PROG, PART) was not really a topic in generative theory. But nominal reference (as opposed to 'verbal' including 'temporal' reference) has been very well studied, via the binding theory and/or more generally indexation procedures. So, one of the possibilities to give body to verbal and temporal interpretation is to build parallels with referentiality.

This track is the one taken by Hamida Demirdache (Chapter 4) in this volume. She explores the temporal construals of perfective vs. imperfective aspect in Sequence of Tense contexts in Spanish and French, in particular under ellipsis. The distribution of past-shifted vs. simultaneous, as well as sloppy vs. strict, temporal construals are taken to support extending to viewpoint aspect a referential approach to tense. She derives the distribution of simultaneous vs. past-shifted readings by extending the analysis of imperfective vs. perfective to embedded contexts. The intricate distribution of strict vs. sloppy readings is derived, she claims, by extending the LF-parallelism constraint on ellipsis to temporal anaphora under ellipsis, specifically the assumption that structural parallelism yields sloppy readings, while referential parallelism yields strict readings.

In relation to the standard view that lots of (if not all) variation is connected to variation in morphology / morpho-syntax, it is to be expected that special situations may lead to specific interpretational problems. An issue of this kind is found by Fernanda Pratas (Chapter 12) in her field work on Capeverdean, a Portuguese-related language spoken in Cabo Verde. According to her, recent descriptions have argued that what seem past tense markers in Capeverdean are allomorphs of a temporal agreement morpheme. She notes that it is true that both *-ba*, and the related (and more complex) form *tava*, are sometimes associated with a past tense (the Topic Time is located before the Time of Utterance) and that this is the case in (i) past progressives and past habituais. On the other hand, they also appear in (ii) subordinate clauses where no past interpretation is certain, such as some conditionals and other modal contexts. Her proposed approach identifies a common point between (i) and (ii): all these structures denote situations with a low degree of accessibility from the speaker's perspective. This (in)accessibility is perceived either in terms of time (as in the first case: one cannot go back to the past) or because external factors may (have) provide(d) an unpredictable outcome (as in the second case). The apparent mismatch is interesting and needs further scrutiny.

The paper by Purificação Silvano, António Leal and João Cordeiro (Chapter 14) bears on adverbial perfect participial clauses in Portuguese varieties and British English. The purpose of this paper is two-fold: (i) to discuss the existing proposal of distinguishing two types of adverbial perfect participial clauses (APC) in European Portuguese; and (ii) to ascertain the key factors behind their temporal interpretation. The authors compare and contrast patterns displayed by APC in European, Brazilian, Mozambican and Angolan Portuguese and British English across a corpus built from newspapers. Their research reveals that the data does not reflect the bipartite division argued before and that, for Portuguese varieties, the position of APC in the sentence and the combinations of some aspectual classes are important to infer temporal relations, whereas, for British English, the anterior orientation/perfect aspect of the perfect participle imposes for most cases an anteriority temporal relation, surpassing the influence of any other factor.

The study of L2 acquisition too is subject to the same approach: morphosyntactic features are central, but examining the actual process is interesting and might also actually bear on the acquisition of the mother tongue. In their paper on infinitival complements, Aida Cardoso, Inês Duarte and Ana Lúcia Santos (Chapter 2) report experimental data from L2 acquisition of European Portuguese. While assuming the Feature Reassembly Hypothesis (FRH), they argue that the acquisition of Exceptional Case Marking (ECM), Inflected Infinitive Structures (IIS) and Prepositional Infinitival Structures (PIC) by Spanish learners of European Portuguese (EP) presents different challenges. Two Acceptability Judgment Tasks show that identifying and reconfiguring the specific features associated with the PIC is difficult and that Spanish native speakers perform better in the case of ECM, whose properties can, for the most part, be transferred from the L1. However, the absence of overt morphosyntactic counterpart for features related to DOM in EP represents the challenge. Finally, the results for the IIS force the authors to question the usual descriptions of the native grammar.

In a similar way, an assumption that L2 acquisition is relatively simple and straightforward with languages that have the same constructions (positive transfer) might sometimes be naive. Sílvia Perpiñán and Rafael Marín (Chapter 10) study aspect in the acquisition of the Spanish locative paradigm. They investigate the development of the expression of this paradigm in the L2 Spanish of Italian-speaking learners. The paper focuses on the mapping of two aspectual features, temporal boundedness and dynamicity, onto the selection of the copular verbs *ser* and *estar* in Spanish locative constructions. In particular, they investigate (i) whether the developmental stages proposed for English-speaking learners in the acquisition of *ser* and *estar* in L2 Spanish reported elsewhere also hold for Italian-speaking learners; and (ii) whether Italian, a language that partially overlaps with the distribution of the Spanish copulas in locative contexts, has a facilitative role or, on the contrary,

has a negative effect on the acquisition of the copulas. Their experimental results indicate that Italian speakers do not present a delay in the acquisition of *estar*, but instead, overproduce it in locative contexts from very early on. They argue that this overproduction of *estar* is due to the readily available mapping of temporal boundedness with *estar* in the grammar of these L2 learners, whereas the presence of the feature dynamicity comes later in L2 development.

Although it is certainly common for linguists studying acquisition to make use of experimental designs, it is also possible to experiment with a psycholinguistic goal. In her paper, Valentina Brunetto (Chapter 1) discusses the processing of clitic pronouns that are not co-arguments. Given an established asymmetry between simple and Exceptional Case Marking predicates in the acquisition of binding Principle B, the study asks whether the notion of coargumenthood plays a role during the on-line processing of clitic pronouns by Italian-speaking adults. She reports experimental evidence from a self-paced reading study and suggests that the time course of pronoun resolution is affected by coargumenthood. In Exceptional Case Marking predicates, comprehenders appear to temporarily consider a feature-matching local antecedent as soon as the clitic trace is processed in its thematic position.

There is a particular problem for most formal approaches to linguistics and that problem is speaker internal variation. It has been noted in several ways; in informal situations the actual use of a language may have different features from a public or formal situation. In the pub you don't speak the same way as with your grandparents and you are more careful when your speech is recorded for television. At the level of the lexicon, it is clear that true synonymy (exactly same meaning in exactly same usages in exactly the same environments) does not exist, because there always will be some difference somewhere. Does the speaker possess two competences, or does the speaker just have one competence with a choice some point? Should a linguist model this as two separate competences or it is one competence with some options at some moment? The issue plays a role with speaker internal variation, but actually it is much more known from the L2 acquisition literature (the statuses of interference types and the notion of interlanguage) and studies of natural bilingualism or dialect speakers going to primary schools. It is really not clear whether in the mind of bilinguals there are actually two separate linguistic competences. It might very well be one somewhat larger competence that at some points builds in a bifurcation / choice without the linguist having to claim that the competences are separate such that this model will just be one model covering various paths. Exactly this discussion forms the background for the paper by Luis López.

In his paper on the I-language of a bilingual, Luis López (Chapter 7) discusses contact phenomena. Against the common-sense notion that bilinguals have two grammatical systems he argues that the linguistic system of a bilingual person should be integrated. In particular, he argues that both the lexicon and the

post-syntactic operations that lead to the externalization systems are integrated. He further argues that the distinction between code-switching and borrowing is spurious and extends the integrated hypothesis to syntactic transfer. His model makes use of Distributed Morphology assumptions to formally describe how an integrated system may work. And indeed this model might be the one that can be motivated by such an approach, suggesting once again that morphology is a central key for linguistic modeling.

In fact, the issue addressed in the phonology paper in this volume shares a property with the paper by López. Although it is not about bilinguals per se, it does bear on two vocabularies with different properties that are united in one phonological system. Loanwords are analyzed as the integration of words of an L2 (or neighboring close L1) into the one phonological system of the L1 with the rather logical consequence that they might give rise to potential OT ranking problems. In terms of López, these speakers using 'regular' L1 words as well as loanwords do not have two separate phonological systems, they have an intrinsically interacting one.

Clàudia Pons-Moll and Francesc Torres-Tamarit (Chapter 11) studied Catalan nativization patterns in the light of Weighted Scalar Constraints. They analyze, from an experimental and formal perspective, the interaction and the implicational relationships between vowel reduction and word-final nasal deletion in Catalan loanwords. They present the results of both a production and a perception test carried out with 31 speakers from the Barcelona area. Loanwords susceptible to undergoing both nasal deletion and vowel reduction display different patterns. The most common solution is underapplication of both processes, followed closely by underapplication of nasal deletion alone and at a large distance by the application of both processes. Finally, underapplication of vowel reduction and application of nasal deletion is unattested, that is, it is a very unlikely nativization pattern. The typology of possible nativizations and the implicational relationships between the processes under scrutiny are analyzed in the framework of Harmonic Grammar under Weighted Scalar Constraints.

The selection of papers in this volume reflects the range of topics and the scope of the approaches that characterizes the circuit of Going Romance and the areas of discussion that we hope to contribute to. We wish to thank all reviewers for judging the 100+ abstracts we received in Summer 2018. We also wish to thank our colleagues associated with the special sessions, for their contribution to the choice of actual scholars who were invited to submit for this volume (we proceeded to a consistent preselection for papers presented at the workshops or at the poster sessions). We finally thank around 50 colleagues for reviewing in a critical way the submitted papers.

Processing clitic pronouns outside coargumenthood

Valentina Brunetto

University of Leeds

Online adult processing of English pronouns is subject to early structural constraints. However, if a sentence fails to provide a licit antecedent for a pronoun, ungrammatical antecedents may be fleetingly considered, causing processing disruption. This paper investigates whether illicit antecedents exert any interference in the processing of clitic pronouns. Given an established asymmetry between simple and Exceptional Case Marking predicates in the acquisition of binding Principle B (Baaauw & Cuetos 2003), this study asks whether the notion of coargumenthood plays a role during the online processing of clitic pronouns by Italian-speaking adults. I report experimental evidence from a self-paced reading study suggesting that the time course of pronoun resolution is affected by coargumenthood. In Exceptional Case Marking predicates, comprehenders appear to temporarily consider a feature-matching local antecedent as soon as the clitic trace is processed in its thematic position.

Keywords: sentence processing, anaphor resolution, Italian clitics, binding, principle B, exceptional case marking, clitic climbing, coargumenthood

1. Introduction

A compelling question in language acquisition research is whether apparently delayed linguistic knowledge reflects delayed maturation of certain grammatical principles or immature language processing abilities. As the evidence from language processing research accumulates, it becomes useful to compare the findings in these two subfields to better understand the role of the language processing system in the development of grammatical knowledge.

In the domain of forward anaphora, the question is how the principles of the Binding Theory, which constrains the encoding of anaphoric dependencies between anaphors and their linguistic antecedents, are implemented during language

comprehension. In brief, Principle A forces reflexives to look for clause-mate structurally available antecedents, whereas Principle B bars clause-mate antecedents for pronouns.

- (1) *John_i says [that Paul_j hurt himself_{j/i}]_i*
 (2) *John_i says [that Paul_j hurt him_{*j/i}]_i*

An important finding is that, both in language acquisition and in language processing, there are patterns of selective ‘difficulty’ (Phillips & Ehrenhofer 2015). Acquisition research reports an asymmetry between the early mastery of Principle A and a ‘delay’ in the implementation of Principle B for strong pronouns (Chien & Wexler 1990), but not for clitic pronouns (McKee 1992). Subsequent research, however, has found patterns of selective difficulty for clitics in certain syntactic structures, such as Exceptional Case Marking (Baauw & Cuetos 2003). In language processing research, a more limited range of languages and linguistic structures has thus far been explored. For English anaphora, the evidence suggests that Principle A is implemented accurately during the processing of reflexives (Clifton, Kennison & Albrecht 1997; Sturt 2003; Dillon et al. 2013), but that the processing of pronouns is more vulnerable to interference effects (Badecker & Straub 2002). However, very little is known about adults’ sensitivity to structurally licit and illicit antecedents during the processing of clitic pronouns, and about the role of locality in structures where the pronoun and its antecedent belong to different predicates. The aim of this study is to investigate whether the selective patterns which have been found in the acquisition of clitics (Baauw & Cuetos 2003) can also be detected in real-time adult language comprehension.

2. Clitics and Principle B: Acquisition evidence and linguistic theory

Thirty years of research on the acquisition of Principle B with clitic pronouns have led to two important discoveries: first, Principle B is acquired early. That is, the developmental ‘delay of Principle B effect’ (Chien & Wexler 1990) is not universal: children acquiring Romance languages perform at ceiling (above 90%) in the same contexts where children acquiring languages with strong pronouns perform at much more variable rates (reported as 50% accuracy on average, as low as 18% in McKee (1992)).

- (3) *Lorso_i lo_{*i} lava* (Italian; McKee 1992)
 The bear him.CL washes
 ‘The bear washes him’

Moreover, pronoun interpretation is influenced by syntactic structure. A key finding is that in complex structures such as Exceptional Case Marking (ECM), where the pronoun is interpreted as an argument of the embedded non-finite verb, difficulties with Principle B are exacerbated. In languages with strong pronouns, children have been found to display a preference for the intrasentential antecedent (below chance performance) in these structures (Philip & Coopmans 1996; Baauw & Cuetos 2003). In languages with clitics, ECM (4) and ‘Faire-Par’ causatives (5) are the only structures where Principle B violations have been attested (Baauw, Escobar & Philip 1997; Brunetto 2015).

(4) *La pecora_i la_{*i} vede [t saltare]*

The sheep her.CL sees jump

“The sheep sees her jump”

(5) *La mamma_i la_{*i} fa [abbracciare t dalla nonna]*

The mother her.CL makes hug by.the grandma

“The mother has her hugged by the grandma”

The discovery that clitics are ‘exempt’ from Principle B errors in child grammar has prompted a large body of theoretical research on binding and coreference and the division of labour between syntax and pragmatics (Grodzinsky & Reinhart 1993; Reinhart 2006). The general consensus is that the interpretation of clitics is regulated by Principle B only (and not by a computation at the syntax-pragmatics interface) due to both their referentially deficient status and the fact that they undergo syntactic movement (Baauw & Cuetos 2003; Brunetto 2015).

However, the existence of Principle B violations in non-coargument contexts such as (4) and (5) raises a theoretical problem. Alternative binding theories which posit a predicate-based notion of locality, such as Reinhart & Reuland (1993), can accommodate the selective fragility of clitics in ECM structures. In this approach Principle B disallows covaluation of two arguments of the same predicate (coarguments) if the predicate is not reflexive-marked. In the ECM construction in (4), however, the pronoun and its antecedent belong to different predicates (non-coarguments). What blocks covaluation in this case is a distinct structural principle on the well-formedness of A-chains. This principle captures the fact that the tail of an A-chain such as [NP ... anaphor] cannot be referential and case marked. This can be seen in a language like Dutch, which allows a simplex anaphor (*zich*) to appear in an ECM structure, but not a pronoun (*haar*):

(6) *Mary_i hoorde [zich_i / *haar_i zingen]*

Mary heard SE / *her sing

“Mary heard herself sing”

The reason a pronoun cannot form an A-chain with the matrix subject is that Agree would copy and overwrite the features from *Mary* to the pronoun *haar* in (6), violating a ‘Principle of Recoverability of Deletions’ (Reuland 2017). In contrast, *zich* is underspecified for phi-features. This allows a syntactic chain to be construed under feature identity without deleting interpretive content.

Given that children’s non-adultlike interpretations of clitics are limited to non-coargument structures, it is important to determine whether the same patterns of selective difficulty can be found in adult language processing. This issue has been recently highlighted by Phillips & Ehrenhofer (2015), who argue that the convergence of language acquisition and processing evidence in many grammatical domains (including anaphora) should be taken as evidence that children’s non-adultlike interpretations stem from the fragility of the developing parser rather than incomplete grammatical knowledge. This claim draws on two appealing ideas: (i) interpretations children don’t entertain are simply interpretations that the parser does not generate (universal constraints), and (ii) errors that children entertain at some developmental stage are interpretations that the adult parser may temporarily entertain during real-time language comprehension. Simply put, adults are better than children at revising and inhibiting illicit interpretations as sentences unfold (see also Snedeker 2013). Phillips & Ehrenhofer (2015) argue that this can explain the special status of Principle B in both acquisition (delay) and processing (selective intrusion profiles). In order to evaluate this proposal, more processing data is needed to address the complexity of Principle B profiles across languages and syntactic structures.

3. Forward anaphora and locality effects in adult processing

Processing research on Principle B focuses on the cues that guide antecedent retrieval during real-time language comprehension. Interpreting an anaphor involves a retrospective search for an antecedent, probing an item in memory which matches a set of retrieval cues while discarding irrelevant items. An obvious cue is agreement between anaphor and antecedent, intended as matching of phi-features. Principle B, as a structural constraint, should also be counted as a cue in antecedent retrieval, because it allows the parser to discard impossible antecedents. The question is whether Principle B is used in immediate retrieval or whether it only applies later, as a ‘filter’ on the result of the retrieval.

There is robust evidence that readers exploit structural cues early on. Adult processing studies employ a gender mismatch paradigm manipulating the phi-features of two potential antecedents. One is structurally inaccessible (local) and one is

accessible (non-local). Early studies argued that the initial set of candidate antecedents includes only entities that are structurally accessible (Clifton et al. 1997), but there has been growing evidence showing that local antecedents interfere with retrieval during processing (Badecker & Straub 2002; Kennison 2003; Chow, Lewis & Phillips 2014). This ‘inhibitory interference’ effect typically occurs later. It results in an interaction between local and non-local antecedent: when the non-local antecedent mismatches the pronoun (*Mary* in (7a, b)), longer reading times may be observed if the local antecedent matches the pronoun (as in (7a)).

- (7) a. *Mary thought [that **Bill** liked **him** a lot]* (→ Inhibitory interference)
 b. *Mary thought [that **Ann** liked **him** a lot]*

This is usually interpreted as a reanalysis/repair process, which starts when the parser has exhausted the search for a licit antecedent. Sturt (2003) called it a ‘de-feasible filter’ suggesting that the binding principles constrain the initial search for an antecedent but may later be violated. Note that this is not per se evidence that antecedent retrieval is unconstrained (in the sense that non-local and local antecedents are being considered at the same time), but only that it takes longer to terminate a search if there is a feature-matching but structurally unavailable (local) candidate. The only evidence that both candidates compete for retrieval would be an interference effect in a sentence like (8a). Badecker & Straub (2002) found longer reading times when both the non-local and the local antecedent matched the pronoun (8a vs. 8b), and suggested that both candidates may be competing during the online processing of anaphoric dependencies. However, this ‘multiple match’ effect has failed to be replicated ever since (Chow et al. 2014).

- (8) a. *John thought [that **Bill** liked **him** a lot]* (Multiple match)
 b. *John thought [that **Ann** liked **him** a lot]*

Less is known about the role of locality and syntactic structure in the online implementation of the binding principles. The only existing evidence that different locality conditions affect the time-course of anaphora resolution comes from studies which compared the interpretation of reflexives in coargumenthood and in picture noun phrases, where anaphors are also considered to be ‘exempt’ (Reinhart & Reuland 1993). In an eye-tracking study, Cunnings & Sturt (2014) found a strong Principle A effect in coargument contexts (9a), with processing disruption detected early (already at the reflexive region in first pass reading) when the stereotypical gender of the local antecedent mismatched the reflexive. In contrast, when the reflexive was embedded inside a picture noun phrase (9b), the local antecedent gender mismatch effect emerged later in the sentence, in the spillover and pre-final region and in second pass reading times.

- (9) a. *Jennifer remembered [that the soldier had positioned herself ...]*
 b. *Jennifer remembered [that the soldier had a picture of herself ...]*

This ‘delayed’ Principle A effect is indicative of a different time course of anaphor resolution outside coargumenthood. Although it is tempting to draw a parallel with the non-coargument contexts discussed here, it should be noted that picture noun phrases are environments where even off-line judgments are mixed (see also Runner, Sussman & Tanenhaus 2003), whereas Principle B violations in ECM constructions are unambiguously ungrammatical in the adult grammar. A structural constraint (whether it is Principle B or simply the A-chain condition, as in Reinhart & Reuland’s binding theory) still needs to be assumed to rule out covaluation between a pronoun and a local antecedent in an ECM structure.

4. Research questions and hypotheses

Summarising so far, studies that have investigated the online processing of strong pronouns indicate immediate sensitivity to structural constraints (adherence to Principle B), but also fleeting interference effects caused by feature-matching local antecedents (typically in later processing, and only as ‘reanalysis’ strategies). This pattern suggests, at least for English, a convergence of developmental data and acquisition data. Based on this premise, I want to investigate the time-course of clitic pronoun resolution in Italian, in both coargument and non-coargument (ECM) contexts. My research questions are therefore the following:

- a. Are there interference effects in the processing of clitic pronouns?
 b. Does (non)-coargumenthood influence the time-course of pronoun resolution?

If Principle B applies to clitics as an ‘indefeasible filter’, we should only observe effects of structurally accessible antecedents, with slowdowns in sentences that fail to provide a matching, non-local antecedent for the pronoun. This is known as a ‘grammaticality effect’. Given that clitics by virtue of their referential deficiency always require a linguistic antecedent, such a grammaticality effect is expected to appear early during processing. Alternatively, I could also observe interference effects caused by the availability of feature-matching local antecedents. If ‘impossible errors’ in acquisition are ‘impossible parses’, one should not expect interference effects in the processing of clitics in coargument contexts, since these are not attested at any developmental stage.

On the other hand, there is less consensus about the status of Principle B in non-coargument contexts such as ECM predicates. If clitics are exempt from Principle B outside coargumenthood, then one could expect a different time course

of clitic processing in ECM structures. Additionally, if ‘possible errors’ in acquisition are ‘temporary misparses’ in adult processing, then we should observe a fleeting interference effect caused by the availability of a feature-matching local antecedent. This interference effect could result in faster reading times (a ‘facilitatory interference’) or slower reading times (‘inhibitory interference’). A facilitatory interference should occur if agreement cues guide antecedent retrieval before structural cues are applied (be it Principle B or the A-chain condition), causing shorter reading times for sentences with feature-matching local antecedents; this, however, has never been found in the processing of anaphora (Dillon et al. 2013). An inhibitory interference, in contrast, should be expected if structural constraints are ‘defeasible’, initially restricting the search to accessible candidates but later allowing the parser to consider illicit antecedents.

5. The experiment

I conducted a self-paced reading study using a within-subject design to compare the processing of clitic pronouns in coargument and non-coargument contexts. All test sentences involved clitic climbing (see Table 1) to ensure that the two sentence types were matched in complexity. Coargument contexts involved modals which, unlike perception verbs, do not assign a subject theta role. Therefore in the modal sentences the subject and the object clitic are coarguments. In contrast, in the ECM sentences the clitic is interpreted as the external argument of the infinitive verb.

Some additional considerations are necessary when analysing the incremental processing of clitic climbing structures. Given that clitics appear preverbally, it is not possible for the parser to predict the argument structure of the predicate (and therefore the structural position of clitic trace) immediately at the clitic region. Only after the infinitive word has been parsed can the pronoun be construed as a coargument (modal sentences) or non-coargument (ECM sentences). The infinitive was therefore the critical region of interest. The other regions of interests are the finite verb (modal vs. perception verb) and the two post-infinitive regions.

5.1 Materials

The experiment included a total of 32 experimental items and 57 fillers. Half of the experimental sentences contained a modal (*volere* = “want”, *potere* = “can”) or perception verb (*vedere* = “see”, *sentire* = “hear”) followed by an infinitive. The fillers also included complex predicates, with more modals (*dovere* = “must”, *sapere* = “know”, *volere*, *potere*) either in a transitive construction with a full NP object

or with a reflexive clitic.¹ The antecedents were always proper names, unambiguous in gender, and both masculine and feminine clitics were used. The average number of syllables in each word region was the same in the two conditions.

The design was 2×2×2 crossing three factors: NONLOCAL SUBJECT (match/mismatch), LOCAL SUBJECT (match/mismatch) and PREDICATE TYPE (coargumenthood vs. non-coargumenthood). Four counterbalancing lists were constructed such that each participant saw one of the four possible factorial combinations for each item. An example of test items is given in Table 1.

Table 1. Examples of experimental items

Condition	Coargumentood (modals)	Non-coargumenthood (ECM)
NONLOCAL MATCH	Emilia ha sognato [che Sandro la voleva licenziare dopo gli scioperi ...] <i>Claudia dreamed [that Sandro her.CL wanted (to) sack after the strikes...]</i>	Paola sapeva [che Claudio la vide prelevare il denaro alla stazione...] <i>Gabriella knew [that Claudia her.CL saw withdraw the money at the station...]</i>
LOCAL MATCH	Emilio ha sognato [che Sandra la voleva licenziare dopo gli scioperi ...]	Paolo sapeva [che Claudia la vide prelevare il denaro alla stazione...]
MULTIPLE MATCH	Emilia ha sognato [che Sandra la voleva licenziare dopo gli scioperi ...]	Paola sapeva [che Claudia la vide prelevare il denaro alla stazione...]
NO MATCH	Emilio ha sognato [che Sandro la voleva licenziare dopo gli scioperi ...]	Paolo sapeva [che Claudio la vide prelevare il denaro alla stazione...]

5.2 Subjects and procedure

Thirty-three Italian university students (between the ages of 19–34, mean 24) were tested in individual sessions. A self-paced word-by-word moving window procedure was used (Just et al. 1982). Sentences initially appeared masked by dashes and participants had to press the spacebar to unmask each word. As one word appeared, the previous one reverted to dashes. Reading times were recorded for each region. The test was preceded by a practice session of six sentences. Comprehension questions appeared after one third of the sentences to ensure that participants attended to the meaning. Participants were instructed to read at a natural pace.

1. An example of a filler sentence is given in (1):

(1) *Raffaele dice che Oriana si vuole iscrivere a una scuola di taglio e cucito.*
“Raffaele says that Oriana herself.CL wants to enrol on a crafts course.”

6. Results

Comprehension questions were answered correctly 87% of the time, indicating that participants attended to the task. Reaction times (RTs) more extreme than 3 times the interquartile range above the 75th percentile or below the 25th percentile by region and condition were classified as outliers, resulting in a loss of 3.6% of the data. Figure 1 displays mean reading times at each presentation region for the embedded clause containing the local subject and the clitic pronoun.

Analyses were conducted on the accuracy data and log-transformed RTs using a linear mixed effect model with random effects for subjects and items, and fixed effects for NONLOCAL SUBJECT, LOCAL SUBJECT, PREDICATE TYPE and their interactions.

A significant main effect of NONLOCAL SUBJECT emerged marginally at the verb region ($t = -1.95, p = .05$) and persisted at the infinitive region ($t = -4.24, p < .0001$) and at the second post-infinitive region ($t = -2.54, p = .01$). Comprehenders were consistently slower when the subject of the main clause (the structurally licit antecedent) didn't match the clitic. However, the model also revealed a significant interaction between LOCAL SUBJECT and PREDICATE TYPE at the verb region ($t = 3.53, p < .001$), followed by a significant three-way interaction between LOCAL SUBJECT, NONLOCAL SUBJECT and PREDICATE TYPE at the critical infinitive region ($t = 2.38, p = .01$) and a significant interaction between NONLOCAL SUBJECT and PREDICATE TYPE at the first post-infinitive region ($t = -2.05, p = .04$).

To explore these interactions, additional analyses were conducted on each level of PREDICATE TYPE (coargumenthood and non-coargumenthood contexts)

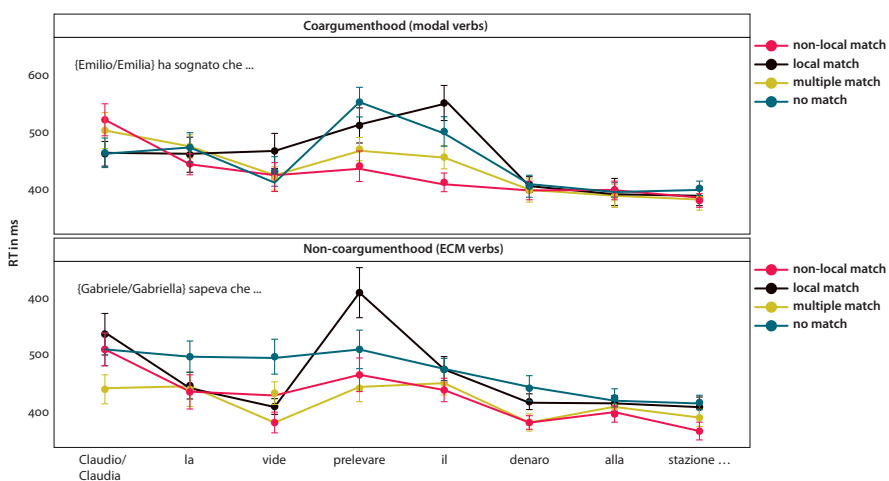


Figure 1. Average self-paced RTs in the coargument condition and in the ECM condition

separately. At the verb region, the effect of LOCAL SUBJECT was significant only in the ECM sentences ($t = -3.53, p < .001$). At the critical infinitive region there was also a significant interaction between NONLOCAL SUBJECT and LOCAL SUBJECT only in ECM contexts ($t = -2.02, p = .04$): reading times were significantly longer in the LOCAL MATCH condition (608 ms) compared to the NO MATCH condition (510 ms). At the first post-infinitive region, the effect of NONLOCAL SUBJECT was significant only in coargumenthood contexts ($t = -4.41, p < .0001$), indicating again that readers were slower when the subject of the main clause mismatched the clitic pronoun.

7. Discussion

In this experiment I asked whether comprehenders temporarily consider structurally illicit antecedents (local subjects) during the processing of clitic pronouns, and whether their (un)willingness to do so changes in different syntactic contexts. To this end, I compared two syntactic structures which differed minimally as to whether the clitic and the embedded subject were arguments of the same predicate or not. The first finding was that a gender mismatch between the clitic and the non-local subject affected reading times in both early processing (at the critical region) and late processing (at the second post-infinitive region). This ‘grammaticality effect’ was modulated by predicate type only at the first region following the infinitive (where it emerged only in coargument contexts). Overall this pattern confirms that agreement cues for non-local and local antecedents are not equally weighted, and indicates that antecedent retrieval is guided by structural constraints regardless of coargumenthood.

Coargumenthood, however, appeared to influence readers’ willingness to temporarily consider local antecedents when a non-local antecedent was not available. In ECM sentences, an interaction between the local and the non-local subject appeared at the infinitive region, resulting in an ‘inhibitory interference’: readers were significantly slower in the LOCAL MATCH condition, in comparison to the NO MATCH condition (difference = 98 ms). This ‘ungrammatical match’ effect is sometimes reported as a spillover effect during the processing of English coargument pronouns (e.g., Kennison 2003; Chow et al. 2014), suggesting a reanalysis process before the search for an antecedent is finally terminated. What sets this result apart from those studies, however, is the fact that this interference effect emerged relatively early in ECM predicates. In fact, the infinitive region was the first point at which the parser could interpret the clitic and the embedded subject as arguments of different predicates. It is important to remember that this interference cannot be taken to suggest a competition between the local and the non-local antecedent. If both candidates were probabilistically considered, then one should expect longer

reading times in the *MULTIPLE MATCH* condition in comparison to the *NON-LOCAL MATCH* condition. This effect was not found. Moreover, the interaction was also accompanied by a main effect of non-local subject (a ‘grammaticality effect’). This reveals an interesting tension: readers differentiated between grammatical and ungrammatical sentences but, additionally, incurred a processing cost as they were inhibiting and discarding the local antecedent in the ungrammatical sentences.

The third major finding of this experiment is that participants never considered illicit antecedents in coargument contexts. Based on this data, the hypothesis that Principle B applies as an ‘indefeasible filter’ in clitic coargument structures appears to be confirmed. In contrast, the processing of non-coargument clitics was susceptible to intrusion from local antecedents, confirming the hypothesis that Principle B is ‘defeasible’ in the ECM construction.

This asymmetry should now be considered in light of the selective fragility which is known to affect non-coargument clitics in acquisition. As Phillips & Ehrenhofer (2015) point out, the existence of the same selective patterns in learners and in online adult parsers has implications for both acquisition and processing theories. While the adequacy of maturational accounts is weakened by the parallel with adult processing, what remains to be explained is what causes certain anaphoric dependencies to be ‘infallible’ and others to be vulnerable to intrusion effects.

Clitics can provide an important piece of this puzzle. It could be argued that coargument clitics satisfy at least one of the properties that appear to characterise ‘infallible’ constraints (Phillips, Wagers & Lau 2011): directionality. Directionality concerns the fact that the parser is better at blocking illicit antecedents during a ‘prospective’ search (Principle C, Strong Cross Over, e.g., Kush, Lidz & Phillips 2017) compared to a retrospective search (Principle B). Clitics, unlike strong pronouns, undergo syntactic movement; hence parsing a clitic involves a ‘prospective’ search for its thematic position. In the bottom-up derivation of coargument clitics, movement creates a local cross-over configuration between the clitic chain and the subject:

- (10) a. $[_{VP} \text{cl}_i [_{VP} \text{subject}_i \text{V } e_i]]$ (coargumenthood)
- ↑
- b. $[\text{subject}_i \text{SEE } [_{\text{Tdef}} [_{VP} \text{cl}_i \text{V } \dots]]]$ (ECM, non-coargumenthood)

In the structures I tested, participants appeared to process coargumenthood as soon as they encountered the predicate, at which point they were immune to interference from the local subject (10a). Perhaps a combination of locality and directionality can explain the special status of non-coargument clitics in ECM. These clitics also involve a ‘prospective’ search, in the sense that the parser has to actively search for the thematic position as soon as the clitic is encountered. However, when this thematic position is identified (10b), the clitic is no longer in a crossover configuration

and the local subject must be discarded by a ‘retrospective’ process. Future research should investigate whether the processing of strong pronouns in ECM predicates gives rise to the same interference effects as those found here, in order to evaluate how different locality conditions, coupled with the requirement for a ‘retrospective’ search for strong pronouns, affect sensitivity to illicit antecedents during online pronoun resolution.

8. Conclusions

In this paper I have shown that coargumenthood influences readers’ sensitivity to Principle B during the online processing of clitics. In coargument contexts, Principle B is implemented with high accuracy, whereas in ECM structures the parser appears to be vulnerable to interference effects. This asymmetry emerged at the point at which the argument structure of the lower predicate could be disambiguated. These findings provide new evidence that the implementation of Principle B in adult online processing is influenced by syntactic structure, giving rise to a pattern of selective difficulty reminiscent of the acquisition profile. If one accepts the claim that errors that are unattested in acquisition are interpretations that are “beyond the generative capacity of the language system” (Phillips & Ehrenhofer 2015: 426), then the task to explain the absence of Principle B violations with coargument clitics falls on processing theories.

Understanding the cause of this pattern of selective fallibility in non-coargument structures requires first of all to understand what makes clitic coargument structures invulnerable to interference effects. At least two properties can be considered: movement and their referentially/pragmatically deficient status. The acquisition literature has primarily focused on the latter (e.g., Grodzinsky & Reinhart 1993) but the interaction between movement (directionality) and locality during adult online processing could also help explain what makes Principle B ‘defeasible’ in some environments but not others. A comparison of clitics and full pronouns in a wider range of syntactic structures should be a question for future investigation.

References

- Baauw, Sergio & Fernando Cuetos. 2003. The interpretation of pronouns in Spanish language acquisition and breakdown: Evidence for the “Principle B Delay” as a non-unitary phenomenon. *Language Acquisition* 11(4). 219–275. https://doi.org/10.1207/s15327817la1104_2
- Baauw, Sergio, Maria A. Escobar & William Philip. 1997. A delay of Principle B effect in Spanish speaking children: The role of lexical feature acquisition. *Proceedings of the Generative Approaches to Language Acquisition Conference, GALA* 97. 16–21.
- Badecker, William & Kathleen Straub. 2002. The processing role of structural constraints on the interpretation of pronouns and anaphors. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 28(4). 748–769.
- Brunetto, Valentina. 2015. The pronoun interpretation problem in Romance complex Predicates. *Lingua: International Review of General Linguistics* 161. 82–100. <https://doi.org/10.1016/j.lingua.2014.09.001>
- Chien, Yu-Chin & Ken Wexler. 1990. Children’s knowledge of locality conditions in binding as evidence for the modularity of syntax and pragmatics. *Language Acquisition* 1(3). 225–295. https://doi.org/10.1207/s15327817la0103_2
- Chow, Wing-Yee, Shewaun Lewis & Colin Phillips. 2014. Immediate sensitivity to structural constraints in pronoun resolution. *Frontiers in Psychology* 5. 630. <https://doi.org/10.3389/fpsyg.2014.00630>
- Clifton, Charles, Shelia Kennison & Jason E. Albrecht. 1997. Reading the words her, his, him: Implications for parsing principles based on frequency and on structure. *Journal of Memory and Language* 36(2). 276–292. <https://doi.org/10.1006/jmla.1996.2499>
- Cunnings, Ian & Patrick Sturt. 2014. Coargumenthood and the processing of reflexives. *Journal of Memory and Language* 75. 117–139. <https://doi.org/10.1016/j.jml.2014.05.006>
- Dillon, Brian, Alan Mishler, Shayne Sloggett & Colin Phillips. 2013. Contrasting intrusion profiles for agreement and anaphora: Experimental and modeling evidence. *Journal of Memory and Language* 69(2). 85–103. <https://doi.org/10.1016/j.jml.2013.04.003>
- Grodzinsky, Yosef & Tanya Reinhart. 1993. The innateness of binding and coreference. *Linguistic Inquiry* 24(1). 69–101.
- Just, Marcel A., Patricia A. Carpenter & Jacqueline D. Wooley. 1982. Paradigms and processes in reading comprehension. *Journal of Experimental Psychology: General* 111. 228–238. <https://doi.org/10.1037/0096-3445.111.2.228>
- Kennison, Shelia. 2003. Comprehending the pronouns her, him, and his: Implications for theories of referential processing. *Journal of Memory and Language* 49(3). 335–352. [https://doi.org/10.1016/S0749-596X\(03\)00071-8](https://doi.org/10.1016/S0749-596X(03)00071-8)
- Kush, Dave, Jeffrey Lidz & Colin Phillips. 2017. Looking forwards and backwards: The real-time processing of strong and weak crossover. *Glossa: A Journal of General Linguistics* 2(1). 70. <https://doi.org/10.5334/gjgl.280>
- McKee, Cecile. 1992. A comparison of pronouns and anaphors in Italian and English acquisition. *Language Acquisition* 2(1). 21–54. https://doi.org/10.1207/s15327817la0201_2
- Philip, William & Peter Coopmans. 1996. The double Dutch delay of Principle B effect. *Proceedings of the 20th Boston University conference on language development, BUCLD 20*, 576–587. Somerville, MA: Cascadilla Press.
- Phillips, Colin & Lara Ehrenhofer. 2015. The role of language processing in language acquisition. *Linguistic Approaches to Bilingualism* 5(4). 409–453. <https://doi.org/10.1075/lab.5.4.01phi>

- Phillips, Colin, Matthew W. Wagers & Ellen F. Lau. 2011. Grammatical illusions and selective fallibility in real-time language comprehension. In Jeffrey Runner (ed.), *Experiments at the interfaces*, Syntax and Semantics, Volume 37, 147–180. Bingley: Emerald Publications. [https://doi.org/10.1108/S0092-4563\(2011\)0000037009](https://doi.org/10.1108/S0092-4563(2011)0000037009)
- Reinhart, Tanya. 2006. *Interface strategies: Optimal and costly computations*, Volume 45. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/3846.001.0001>
- Reinhart, Tanya, & Eric Reuland. 1993. Reflexivity. *Linguistic Inquiry* 24(4). 657–720.
- Reuland, Eric J. 2017. Why is reflexivity so special? Understanding the world of reflexives. *Studia Linguistica* 71(1/2). 12–59. <https://doi.org/10.1111/stul.12070>
- Runner, Jeffrey T., Rachel S. Sussman & Michael K. Tanenhaus. 2003. Assignment of reference to reflexives and pronouns in picture noun phrases: Evidence from eye movements. *Cognition* 89(1). B1–B13. [https://doi.org/10.1016/S0010-0277\(03\)00065-9](https://doi.org/10.1016/S0010-0277(03)00065-9)
- Snedeker, Jesse. 2013. Children's sentence processing. In Roger van Gompel (ed.), *Sentence processing*, 201–232. New York: Psychology Press.
- Sturt, Patrick. 2003. The time-course of the application of binding constraints in reference resolution. *Journal of Memory and Language* 48(3). 542–562. [https://doi.org/10.1016/S0749-596X\(02\)00536-3](https://doi.org/10.1016/S0749-596X(02)00536-3)

Infinitival complement clauses

Data from L2 acquisition of European Portuguese

Aida Cardoso, Inês Duarte and Ana Lúcia Santos

University of Lisbon

Assuming the Feature Reassembly Hypothesis (FRH) (Lardiere 2008, 2009), we argue that the acquisition of Exceptional Case Marking (ECM), Inflected Infinitive Structures (IIS) and Prepositional Infinitival Structures (PIC) by Spanish learners of European Portuguese (EP) presents different challenges. Two Acceptability Judgment Tasks show that identifying and reconfiguring the specific features associated with the PIC is a difficult task and that Spanish native speakers perform better in the case of ECM, whose properties can, for the most part, be transferred from the L1. However, the absence of an overt morphosyntactic counterpart for features related to Differential Object Marking (DOM) in EP represents a challenge. Finally, the results for the IIS present an interesting case study, since they force us to question the usual descriptions of the native grammar.

Keywords: syntax, L2 acquisition, feature reassembly hypothesis, European Portuguese, Peninsular Spanish, infinitival complements, inflected infinitive, prepositional infinitival construction, causative verbs, perception verbs

1. Introduction

This study focuses on the non-native acquisition of infinitival complement clauses of perception (e.g., *ver* “to see”) and causative verbs (e.g., *deixar* “to let”) by Spanish learners of EP. Spanish and EP make available a set of different non-finite complement structures of causative and perception verbs, involving different sets of features associated with different functional projections. Importantly, there is not total correspondence between the structures available in EP and in Spanish, which makes this a topic with interesting potential implications for non-native acquisition. Exceptional Case Marking (ECM) is available in both EP (*a Maria deixou as crianças comer um gelado*. “Maria let the children eat an ice cream.”) and Spanish (*María dejó a los niños comer un helado*. “Maria let the children eat an ice cream.”) with similar syntactic properties, namely, in both languages (similar to what happens in

other languages, such as English), ECM is characterized by an Accusative embedded subject and an uninflected infinitive. On the contrary, the Inflected Infinitival Structure (IIS), a case in which the embedded complement is an inflected infinitive clause with a Nominative subject (*a Maria deixou as crianças comerem um gelado*. “Maria let the children eat an ice cream.”), is only available in EP.¹

While ECM and IIS are available with causative and perception verbs, there are also two distinct structures in EP and in Spanish that are only possible with perception verbs: the Prepositional Infinitival Construction (PIC), in EP (*A Maria viu as crianças a comer um gelado*. “Maria saw the children eating an ice cream.”), and the Gerund Construction (GC), in Spanish (*María vio a los niños comiendo un helado*. “María saw the children eating an ice cream.”). As Casalicchio (2019) notes, these two structures, which convey an aspectual progressive value, share a similar syntax and tend to exist in complementary distribution across the Romance languages. In other words, if a language makes the PIC available, it does not make available the GC in the same context, and vice-versa. This is the case for EP and Spanish.²

Over the next sections, we describe the properties of these structures in more detail, but, for now, it is important to keep in mind that the similarities and differences between the two languages can be thought of in terms of features associated with specific functional categories. For instance, for ECM and the IIS, we can think of their differences in terms of the properties associated with Case (Accusative vs. Nominative subjects) and Number (uninflected vs. inflected infinitives) features in T. As for the PIC and the GC, they both express a non-culminated event (Demirdache & Uribe-Etxebarria 1997) and share a progressive aspectual feature. Nevertheless, in the two structures, this aspectual feature has different morphosyntactic counterparts: in the PIC this feature is encoded in a preposition (*a* “to”) plus an infinitival verb form, whereas in the GC it is encoded in the gerund verb form.

Considering this, we discuss the non-native acquisition of these structures by following the theoretical principles and general predictions of the Feature Reassembly Hypothesis (FRH) (Lardiere 2008, 2009, a.o.). This hypothesis focuses

1. EP and Spanish also have available another infinitival complement of causative and perception verbs, the *faire-inf* construction. Nevertheless, this structure is much less frequent in EP than in Spanish. Furthermore, as noted by Gonçalves (1999), there is a difference in the selection properties of perception verbs regarding *faire-inf* in both languages. In EP, *faire-inf* is only available with perception verbs if the verb in the embedded domain (that is, the infinitival verb form) is an intransitive verb, in which EP differs from Spanish. For these reasons, *faire-inf* will not be discussed in this paper.

2. The GC is also available in EP as a complement of perception verbs, but not in the standard variety. In such context, this structure is only available in southern varieties and in some insular dialects (Pereira 2015).

on the articulation of different grammar components and the role of features, their formal configurations (feature bundles), and lexical representation in the L2 acquisition process. By taking into account the role of features and their morphosyntactic encoding in the non-native acquisition process, this hypothesis allows us to make predictions considering the contrasts between EP and Spanish that have been mentioned thus far. On the contrary, previous hypotheses, such as the Interpretability Hypothesis (Hawkins & Hattori 2006; Tsimpli & Dimitrakopoulou 2007), whose predictions regarding the acquisition process rely on the parametrical properties of the L1 and the L2, do not allow us to account for the role of more fine-grained differences between the L1 and the L2 in the acquisition process. Thus, it is our belief that the topic of this paper can contribute to further discussion of the explanatory power of the FRH. Furthermore, the language pairing EP-Spanish is still understudied under the FRH and the typological proximity of these two languages brings to the table new challenges in terms of non-native acquisition that provide an optimal context to test the capacity of the FRH to make predictions based on subtle differences between the L1 and the L2.

In § 2, we focus on the properties of the infinitival complements under study and in § 3 we lay out the predictions that can be made, under the FRH, about the specific challenges of the acquisition of ECM, the IIS, and the PIC in EP L2 by Spanish native speakers. In §§ 4–5, we present our experimental study, whose main conclusions are summarized in § 6.

2. Infinitival complement clauses in European Portuguese and in Peninsular Spanish

2.1 Exceptional case marking (ECM) and inflected infinitive structure (IIS)

ECM can be described as an infinitival complement with a bare infinitive and an Accusative subject, which receives its Case not from the embedded verb (the infinitive), but from the matrix verb, in this case, the causative or perception verb. The sentences in (1) and (2) illustrate these properties for EP and Spanish with the perception verb *ver* (“to see”), but the same applies to causative verbs.³

3. The literature on Spanish considers that ECM is available with perception and causative verbs, namely *dejar* (“to let”) (Hernanz Carbó 1999), but that with other causatives this structure is available with semantic restrictions (Torrego 1998). This would mean that causative verbs have different selection properties in EP and in Spanish, since EP does not present such restrictions. We take this into account in the design of the experimental tasks (see § 4).

- (1) (EP)
- a. *A Maria viu os meninos comer o bolo.*
the Maria saw the children eat.INF the cake.
“Maria saw the children eat the cake.”
 - b. *A Maria viu-os comer o bolo.*
the Maria saw-CL.ACC eat.INF the cake.
“Maria saw them eat the cake.”
- (2) (Spanish)
- a. *María vio a los niños comer el bizcocho.*
María vio to the children eat.INF the cake
“María saw the children eat the cake.”
 - b. *María les vio comer el bizcocho.*
María CL.DAT saw eat.INF the cake.
“María saw them eat the cake.”

Moreover, ECM shares other syntactic properties in EP and Spanish, for instance, the possibility of reflexive pronouns (3) and of negation markers (4) in the embedded domain, and the impossibility of clitic climbing (5) (Gonçalves 1999 et seq., for EP; Bordelois 1988; Treviño 1994; Hernanz Carbó 1999, for Spanish). The examples below are from EP, but the same holds for Spanish.

- (3) *A mãe deixou as meninas pentear-se.*
the mother let the girls comb.INF.REFL
“The mother let the girls comb their hair.”
- (4) *O professor deixou os alunos não fazer o trabalho de casa.*
the teacher let the students not do the homework
“The teacher let the students not to do the homework.”
- (5) a. *A Maria viu os meninos comer o bolo*
the Maria saw the children eat.INF the cake.
“Maria saw the children eat the cake.”
- b. **A Maria viu-o os meninos comer.*
The Maria saw.CL.ACC the children eat.INF
“Maria saw the children eat it.”
- c. *A Maria viu os meninos comê-lo.*
the Maria saw the children eat.INF.CL.ACC
“Maria saw the children eat it.”

However, EP and Spanish differ in that constituents in an Object position in Spanish are subject to DOM and, consequently, (a subset of) Accusative constituents in Spanish are preceded by *a* (“to”), presenting a morphological Dative marking. In EP, however, there is no DOM and, therefore, the embedded subject surfaces as

an Accusative constituent with no additional marking (1). Traditionally, the constituents subject to DOM have been analysed as Accusatives with a Dative morphological marking associated with semantic features such as [+animated] and [+specific] (Brugé & Brugger 1996; Torrego 1998; Rodríguez-Mondoñedo 2008, a.o.). Following Longobardi's (2005) definition of feature, according to which a feature is grammaticalized if it "must obligatorily occur and be valued in a certain structure ..." (Longobardi 2005: 410), we can argue that, in Spanish, the semantic features [animated] and [specific] are functional features grammaticalized when associated with an Accusative constituent, in which Spanish differs from EP. As we will see later, the issue of DOM is also relevant in the acquisition of the PIC.

In addition to the ECM structure, EP grammar also includes inflected infinitives, and therefore the IIS (6).

- (6) a. *A Maria viu os meninos comerem o bolo.*
 the Maria saw the children eat.INF.3PL the cake.
 "Maria saw the children eat the cake."
 b. *A Maria viu eles comerem o bolo.*
 the Maria saw they.NOM eat.INF.3PL the cake
 "Maria saw them eat the cake."

Although IIS behaves as ECM regarding the distribution of clitic pronouns, reflexive pronouns, and negation markers (Gonçalves 1999), the two structures crucially differ in the Case licensing of the embedded subject, as expected. As shown in (6b), IIS is characterized by a Nominative subject occurring with a verb form with overt morphological subject-verb agreement in person and number (the inflected infinitive). Despite the fact that this structure is described in the literature as part of standard EP, recent studies have provided data that raise questions about the status of the IIS as a complement of perception or causative verbs in the EP grammar. Barbosa, Flores & Pereira (2018) tested the acceptability of this structure (among others) by monolingual and heritage EP speakers. It is also important to mention that the target sentences used in the design of the AJT in Barbosa et al. (2018) presented a Nominative pronominal subject (similar to the sentence in (6b)), in order to avoid any ambiguity with other possible infinitival types of complement structures. The results show a very low (4.1%) acceptance rate of the IIS in this context and with pronominal subjects by monolingual EP speakers.

Previous studies on infinitival complements of causative and perception verbs in EP have raised theoretical issues about different defectiveness levels associated with these structures.⁴ Under the Minimalist framework, Gonçalves (1999); Gonçalves

4. For an earlier analysis, see also works by Raposo (1981, 1987).

& Duarte (2001), and Duarte & Gonçalves (2002) proposed a comparative analysis of ECM and IIS⁵ in terms of a functional defectiveness scale. The authors argue that ECM has a more defective functional structure than IIS, corresponding to a defective TP: on the one hand, it allows for negation markers and clitic pronouns in the embedded domain, evidence for a T projection; on the other hand, the licensing of an Accusative subject by the matrix verb reflects the incomplete nature of T.

In a more recent proposal, Hornstein, Martins & Nunes (2008) also discuss the defectiveness of T in ECM, but the focus moves from the functional structure as a whole to the specific feature properties of T. In fact, the authors argue that T is defective in ECM because it lacks a Person feature. This proposal is motivated by the fact that the embedded subject has its Case valued by an external probe (the matrix verb) when the presence of a Person feature would imply the existence of a probe in the embedded domain able to attribute Nominative Case to the subject (in line with Chomsky (2001)).

As for IIS, the fact that there is subject-verb agreement and that a Nominative subject is licensed by the embedded verb leads Gonçalves (1999 and subsequent works) to argue in favour of an AgrS projection. Considering more recent proposals under the Minimalist framework, such as Chomsky's (2008) definition of phases, we can hypothesise that the IIS may correspond to a non-defective T, if we assume that it aggregates the properties of AgrS, or to a CP projection. The latter hypothesis would also be in line with other works on inflected infinitives in different contexts than the one discussed here; such works have argued that Inflected Infinitives are CP projections (e.g., inflected infinitive clauses as subjects, Adjuncts, and Complements of factive, declarative or epistemic verbs).⁶

For Spanish, ECM has been studied mainly considering perception verbs (Contreras 1987; Castillo 2001).⁷ Castillo (2001) departs from the analyses of ECM as a defective TP and argues that it is a Small Clause with no TP. More specifically, ECM would be a VP embedded in multiple AspP nodes with different semantic features. The author focuses on the simultaneous nature of the events described in the matrix and in the complement clauses, assuming that some properties would derive from this analysis, namely the impossibility of negation markers and compound tenses (the perfective *have*) to occur in the embedded domain.

5. The authors also include *faire-inf* in their analysis, arguing that this is the most defective structure, followed by ECM and IIS.

6. For discussion of other Inflected Infinitive Structures in EP, see Raposo (1987); Ambar (1992); Madeira (1994); Duarte, Gonçalves & Miguel (2005), and Gonçalves, Santos & Duarte (2014).

7. In the case of causative verbs, the literature has focused on the study of complex predicate structures (*faire-inf*), mainly with the verb *hacer* "to make" (e.g., Bordelois 1988; Treviño 1994; Torrego 2010; Tubino 2010).

Nevertheless, we can argue that in Spanish, as in EP, the restrictions to the occurrence of negation in the embedded domain are motivated by semantic and pragmatic constraints imposed by the specificities of (direct) perception verbs. On the one hand, these restrictions do not apply to causative verbs (7) and not even to perception verbs given an adequate pragmatic context, in which an event is still susceptible to be perceived (8).

- (7) *La madre dejó a los niños no comer la sopa.*
 the mother let to the children not to eat the soup
 “The mother let the children not to eat the soup.”
- (8) *La entrenadora vio a los atletas no completar la corrida.*
 the coach saw to the athletes not finish the race
 “The coach saw the athletes not finish the race.”

Regarding the impossibility of occurrence of the past perfect in the embedded domain, it does not seem to depend on the syntactic structure itself, but on the temporal dependence between the matrix and the infinitive verb. Evidence for this can be found in less defective complements in EP, such as the IIS, which do not allow for the past perfect to occur in the embedded domain as well (9). In the case of finite complements, the past perfect is possible but, crucially, not with a direct perception interpretation (10). Given the properties discussed so far, we argue that ECM shares a relevant property in Spanish and EP: a defective T.

- (9) **A Maria viu os meninos terem comido a sopa.*
 the Maria saw the children have.AUX.3PL eaten the soup.
 “Maria saw the children eat the soup.”
- (10) *A Maria viu que os meninos tinham comido a sopa.*
 the Maria saw that the children have.AUX eaten the soup.
 “Maria saw that the children have eaten the soup.”

So far, we have focused on ECM and IIS as these are traditionally described in EP grammars as part of the standard variety (Duarte 2003; Barbosa & Raposo 2013). Nevertheless, an additional structure has been described for EP (and is also not available in Spanish): an ECM structure with what seems to be an inflected infinitive (Hornstein et al. 2008; Barbosa et al. 2018; Martins 2018), as shown in (11).

- (11) *O professor deixou-os terminarem o exame.*
 the teacher let-CL.ACC finish.INF.3PL the exam
 “The teacher let them finish the exam.”

This structure is available in EP, but its acceptability can be subject to some degree of variation. Since there are few studies based on experimental data about this

structure, it is difficult to determine to what extent it is accepted by EP native speakers. An exception to this is the aforementioned study by Barbosa et al. (2018), which presents results showing that this structure achieves high acceptance rates among monolingual EP speakers. In fact, comparing the results for instances of ECM with an inflected infinitive (such as 11) against instances of IIS (such as 6b), both with pronominal embedded subjects, Barbosa et al. (2018) show that ECM with an inflected infinitive (i.e., overt subject-verb agreement morphology) has higher acceptability rates than IIS.

It needs to be noted that it is still not clear in what contexts this structure is accepted by native speakers. According to Hornstein et al. (2008), this structure is only possible with the 3rd person plural, being ungrammatical in the case of any other person specification (see the contrast between (11) and (12)).

- (12) **O professor deixou-nos terminarmos o exame.*
 the teacher let-CL.ACC finish.INF.1PL the exam
 “The teacher let us finish the exam.”

Hornstein et al. (2008) analyse this structure as an infinitive with overt morphology assigned by the morphological component (bearing no relevance in the syntactic computation component, in line with Chomsky (2008), and, thus, not a true inflected infinitive). Hence, the authors argue that the morphological component may assign a default form to the verb, which, in this case, is “a valued number feature with a default person feature, yielding a ‘defective inflected infinitive’” (Hornstein et al. 2008: 14) – the default person would be the 3rd and the valued number feature would correspond to plural marking. Nevertheless, data from Barbosa et al. (2018) show some gradience in the rejection rates of this structure depending on the specification of Person (1st plural, 2nd singular, and 3rd plural): on the one hand, sentences with an Accusative 3rd person plural subject plus an inflected infinitive have the lowest rejection rates by monolingual EP speakers (45.5%), which seems to be in line with Hornstein et al. (2008); on the other hand, the data also reveal an asymmetry between the rejection of sentences with 1st person plural subjects (88%) and the 2nd person singular and 3rd person plural subjects (62.5% and 45.5%, respectively). In the present study, we only test sentences with 3rd person plural Accusative subjects plus Inflected Infinitive, but previous studies make clear that more data are needed in order to clarify the status of this structure in EP grammar. In § 4, when we describe our experiment, we refer to this structure as ‘ECM with inflected infinitive, non-standard’ – the designation is merely descriptive.

2.2 Prepositional infinitival structure (PIC)

Perception verbs in EP select an additional type of infinitival complement, the PIC. This structure, which is exemplified in (13), can be described as presenting an embedded Accusative subject and an infinitive verbal form that is preceded by the preposition *a* “to”. Another specificity of this structure is that it can occur either with an uninflected (13a) or an inflected infinitive (13b), but a Nominative subject is never licensed in this context, even when an inflected infinitive occurs.

- (13) a. *O profesor viu-os a ler a gramática.*
 the teacher saw-CL.ACC to.ASP read.INF the grammar
 “The teacher saw them reading the grammar.”
- b. *O profesor viu-os a lerem a gramática.*
 the teacher saw-CL.ACC to.ASP read.INF.3PL the grammar
 “The teacher saw them reading the grammar.”

The PIC is only available in EP. Spanish has a different non-finite complement available with perception verbs, the Gerund Construction (GC), which shares a set of properties with the PIC. As shown in (14), the GC also presents an embedded Accusative subject (with DOM), but it presents a gerund verb form (instead of an infinitival verb form) and no preposition.

- (14) *Vi a Juan conduciendo una furgoneta blanca.*
 saw.1SG to Juan driving.GER a van white
 “I saw Juan driving a white van.” (Rafel 2000: 202 (44a))

Most importantly, both the PIC and the GC express a progressive aspectual value, therefore obligatorily corresponding to non-culminated events (Dermidache & Uribe-Etxebarria 1997), in which they differ from the non-finite complements of causative and perception verbs discussed before. Furthermore, traditionally both structures have been considered Small Clauses (Raposo 1989; Barbosa & Cochofel 2005; Rafel 2000; Casalicchio 2013, 2019).

Nevertheless, for the analysis of the PIC, we can identify three fundamental studies: Raposo (1989); Duarte (1992), and Barbosa & Cochofel (2005). Raposo (1989) considers the PIC to be a Small Clause headed by the preposition *a* “to”, since (i) it has a distribution similar to other Small Clauses (but crucially different from other infinitival complements, such as the IIS), and (ii) it corresponds to a single constituent. Furthermore, Raposo (1989) argues in favour of PIC being a Control structure in which the subject of the PIC (the subject of the Small Clause) controls an empty category (PRO, with uninflected infinitive, or *pro*, with inflected infinitive) in the infinitival clause. The author also considers the PIC an

ECM instance, since it exhibits transparency effects and allows for the matrix verb to license Accusative Case in the subject of the Small Clause, regardless of the type of infinitive in the embedded domain (inflected or uninflected).

Duarte's (1992) analysis departs from Raposo's (1989), arguing against an analysis of PIC as a Small Clause and a Control structure. Instead, Duarte (1992) highlights the aspectual value denoted by the sequence formed by the preposition and infinitival morpheme (reanalysed as a discontinuous morpheme) and claims that the PIC is an AspP.

Barbosa & Cochofel (2005) incorporate aspects of the previous analyses and propose that the PIC is a Small Clause and a Control structure (in line with Raposo 1989), but consider the preposition *a* "to" to be an aspectual head, following Duarte (1992). Moreover, they associate the progressive value of the PIC with an analysis of the *a* as a preposition of central coincidence (Hale 1986), since these types of prepositions denote a durative or progressive value.

As for the GC, different analyses have also been suggested. Considering the topic of this work, we focus on two comparative studies of PIC and GC provided by Rafel (2000) and Casalicchio (2013, 2019).⁸ Rafel (2000) analyses PIC and GC as Complex Small Clauses, for he considers that these constructions have a more complex syntactic structure than other types of small clauses. Hence, he argues that the PIC corresponds to a PP and that the GC is a functional projection FP, but, crucially, both small clauses select a CP complement.⁹ Furthermore, along the lines of what has been proposed for the PIC by Raposo (1989) and Barbosa & Cochofel (2005), he considers the PIC and the GC to be Control structures.

On the other hand, Casalicchio (2019) proposes a unified analysis of PIC and GC based on the argument that both structures have a preposition *a* "to" in which the progressive aspectual value is encoded. However, in the case of GC, the preposition is null and incorporated in the verb form. Assuming this, Casalicchio proposes that the PIC and the GC are Small Clauses corresponding to FP and that their subject is merged in Spec,vP and then moves to Spec, FP. Therefore, in opposition to Rafel (2000), Casalicchio argues that the PIC and the GC are not Control structures and that they are not CP since, as the author notes, neither modal verbs nor auxiliaries can occur in the embedded domain of both structures. The main difference between the PIC and the GC lies in the position in which the preposition is merged: later (in a higher AspP) in PIC (Casalicchio 2019: 98, (52)) and earlier (in a lower AspP) in GC (Casalicchio 2019: 98, (51)). As a result of this difference, the PIC

8. Rafel (2000) and Casalicchio (2013) also include Pseudo-relatives in their comparative studies, but the discussion of these structures is not the aim of the present study.

9. As Casalicchio (2013, 2019) notes, Rafel's analysis for GC focuses primarily on data from English and not from Spanish or other Romance languages.

has a more expanded syntactic structure, thus predicting that some elements can intervene between the preposition and the verb, such as certain types of adverbs and negation markers, as previously noted by Barbosa & Cochofel (2005).

3. Non-native acquisition of European Portuguese infinitival complement clauses: A case of feature reassembly

Our discussion is framed by the Feature Reassembly Hypothesis (FRH) (Lardiere 2000, 2008, 2009, a.o.). Unlike previous hypotheses that focus on the role of parameters and parameter setting in L2 acquisition, Lardiere argues that to achieve a better understanding of the L2 acquisition process and of the variability that often characterizes the performance of L2 learners, we need to look into more complex processes. These more complex processes involve, according to the author, features and their configuration in feature bundles. This is why Lardiere argues that the central status of parameter resetting must be reconsidered (Lardiere 2008), since the acquisition of binary parametric values does not account for the individual-speaker variability, particularly considering morphosyntactic properties.

It should be noted, though, that the focus of the FRH is not a simple contrastive comparison of what features are available in the L1 and the L2. What is relevant is to observe how features are assembled in syntactic configurations in each language and how those features are encoded from the morphosyntactic and morpholexical point of view. Hence, the FRH moves the theoretical focus from parameter resetting to the process of acquisition of features and their configuration in feature bundles associated with different lexical representations in the L1 and in the L2 (Lardiere 2009). Thus, the FRH brings a more fine-grained view of the process of acquisition by allowing the inclusion of variation in the discussion (Rothman & Slabakova 2018).

Lardiere (2009: 175) summarizes the task of the L2 learners in three main questions: (i) with which functional categories are the relevant features associated in the syntax of the L2 and how this may differ from the L1?; (ii) in which lexical items of the L2 are the relevant features expressed?, and (iii) which forms are optional and which are obligatory and what determines an obligatory context? In the present study, our predictions address mainly the two first questions.

Furthermore, the FRH works on the basis of the assumption that L2 learners already have a system of formal features, assembled into the items in the lexicon (both lexical and functional categories) of the L1, and that different formal configurations of features between the L1 and the L2 may be more difficult for L2 learners. What is more, given that the FRH is a hypothesis that assumes full access to the UG, it also assumes that the acquisition of any property or structure is possible as long as a formal contrast (between the L1 and the L2) is detected.

Lardiere (2008, 2009) operationalizes the acquisition process in two main tasks: mapping and reassembling. Mapping corresponds to the initial stages of acquisition of an L2. In this stage, learners associate (or map) specific L1 feature combinations with the L2 configurations that they perceive as closest to the ones of the L1. In the reassembly stage, we find more complex operations since learners might need to add features that are not part of the inventory of the L1, but also avoid L1 features that are not encoded in the L2. It is also in this stage that learners reassemble features that are not associated with the same configurations and representations in the L1 and in the L2.

We can associate different degrees of complexity with the tasks of mapping and reassembly: a property whose acquisition requires only mapping between the L1 and the L2 entails less complexity than one that requires reassembling. It follows, then, that properties requiring only mapping are easier acquired by L2 learners than the ones requiring also reassembling. More importantly, building on the general predictions of the FRH, Montrul & Yoon (2009) hypothesise that, in the reassembling task, adding features may be easier than avoiding features that are part of the inventory of the L1 but not of the L2. According to the authors, “it is easier to detect the existence of something in the input and infer its availability and correctness than to notice the absence of something and thus infer impossibility of a given interpretation” (Montrul & Yoon 2009: 308).

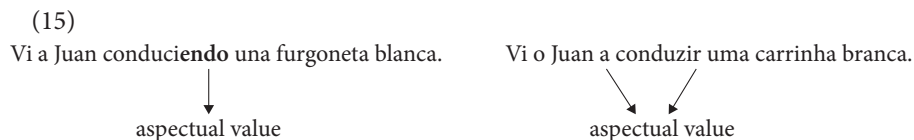
Given the discussion so far, we can now question what it implies for Spanish learners of L2 EP to acquire ECM, IIS, and PIC under the FRH. In the case of ECM, we have shown that EP and Spanish have available an infinitival complement that shares a relevant set of properties and that can be analysed as a defective T regarding its ϕ -features. Considering this, we predict that this structure will be acquired earlier by the L2 EP learners since the relevant feature configuration (the features associated with T) can be mapped from the L1 to the L2.

As for the IIS, this complement is characterized by a Nominative subject licensed by an inflected infinitive and is only available in EP. In § 2.1 of this paper, we hypothesized that the IIS could be analysed as a non-defective T or as a CP. Either way, we assume that acquiring this structure entails acquiring the feature bundle configured in an infinitival T which is non-defective regarding its ϕ -features. This means that to acquire the IIS, Spanish native speakers must reassemble the properties of T by adding features that are not part of the inventory of their L1.¹⁰ Hence, we hypothesise that this structure, whose acquisition entails reassembly by

10. Spanish has Personal infinitives, i.e., structures in which an infinitive licenses a Nominative subject. Nevertheless, these do not have overt subject-verb agreement morphology and they are not available in complement position, but only as adverbial and subject clauses or in impersonal contexts (Mensching 2000).

adding features in feature configurations not available in the L1, will be acquired after the ECM.

Regarding the PIC, we have seen that both EP and Spanish make available a type of complement of perception verbs that expresses a progressive aspectual value and that this complement is generally analysed as a Small Clause. However, the progressive aspectual value has different morphological counterparts in the GC and in the PIC, as shown in the schematic representation in (15).



“I saw Juan driving a white van.”

[sentences adapted from Rafel 2000: 202 (44a)]

Hence, we question whether acquiring the PIC could present a challenge for L1 Spanish speakers due to possible difficulties in identifying the morphosyntactic contrast illustrated in (15), that is, the difficulty for these EP L2 learners may begin in the mapping stage itself. Even taking into account a more unified account of PIC and GC, like the one proposed by Casalicchio (2013, 2019), we can hypothesise, under the FRH, that Spanish speakers could show some degree of difficulty in mapping and reassembling the progressive aspectual value encoded in a preposition incorporated (not spelled-out) in a gerund verb form (GC) to the same aspectual value encoded in a (spelled-out) preposition plus an infinitival verb form (PIC). In other words, mapping and reassembling the [progressive] feature, which has different morphological counterparts in the L1 and in the L2, may be a complex task. An additional difficulty may lie in the infinitival verb form itself, which can be either inflected or uninflected, and whose acquisition, in the case of the inflected infinitive, entails adding person and number features not available in a similar context in the L1.

Finally, we must address DOM and its implications for the acquisition process. As we have seen, in Spanish the Accusative subject of ECM shows DOM and, thus, presents a Dative morphological marking associated with the semantic features [+animated] and [+specific]. For Spanish speakers to successfully acquire the Case marking properties of ECM (and PIC) subjects in EP, they need to be able to avoid L1 features that are not encoded in the L2. In other words, if we take Longobardi’s (2005) definition of ‘feature’ into account, Spanish speakers must acquire the impossibility of grammaticalization of semantic features such as [+animated] and [+specific] in the L2. Since avoiding L1 features not encoded in the L2 may be a complex task (Montrul & Yoon 2009), we predict that Spanish learners

of EP struggle with the reconfiguration needed for acquiring the L2 Case marking properties. This may affect the acquisition of ECM in L2 EP, otherwise a structure that we predict will offer no special difficulties to speakers.

4. Experimental design: Methodology and participants

To test our predictions, we designed two Acceptability Judgment Tasks (AJT), in order to get complementary data on the acquisition of ECM, IIS, and PIC. The first AJT allows us to compare the acquisition process of ECM and of IIS by testing the acquisition of Case and ϕ -features of the embedded domain in each structure. With this goal in mind, we designed an AJT with four conditions: (i) Accusative subject + inflected infinitive (ECM with inflected infinitive, non-standard); (ii) Accusative subject + uninflected infinitive (ECM, i.e., standard ECM); (iii) Nominative subject + inflected infinitive (IIS); (iv) Nominative subject + uninflected infinitive (expected to be judged as ungrammatical). We had a total of 16 test items (4 per condition: 2 with the perception verb *ver* ‘to see’ and 2 with the causative verb *deixar* ‘to let’).¹¹ Table 1 presents the different conditions.¹²

Table 1. Conditions of the AJT testing ECM and IIS

Condition 1: Accusative subject + inflected infinitive

^{OK/*}O Manuel deixou-os / viu-os lerem uma revista.
 the Manuel let-CL.ACC / saw-CL.ACC read.INF.3PL a magazine.
 ‘Manuel let/saw them read a magazine.’

Condition 2: Accusative subject + non-inflected infinitive (ECM)

^{OK}O editor deixou-as / viu-as escrever um artigo.
 the editor let-CL.ACC / saw-CL.ACC write.INF an article.
 ‘The editor let/saw them write an article.’

11. We only included one verb of each class to avoid possible lexical effects that could affect the participants’ answers. The choice of *ver* ‘to see’ and *deixar* ‘to let’ was motivated by the fact that, in terms of ECM, the selection properties of these verbs are similar in EP and in Spanish. The comparison between different causative and perception verbs should be considered in future research.

12. Since Condition 1 corresponds to a non-standard structure, whose acceptability is subject to variability, we decided to signal the sentences in this condition with the notation ‘ok/*’ in Table 1, to indicate that the literature does not make clear predictions on the general grammaticality status of this type of sentences for monolingual EP speakers. What was known at the time of our data collection was that some EP speakers accepted this type of sentences, but others did not.

Table 1. (continued)

Condition 3: Nominative subject + inflected infinitive (IIS)

^{OK}O professor deixou / viu eles tocar piano.
 the teacher let / saw they.NOM play.INF.3PL piano.
 “The teacher let/saw them play the piano.”

Condition 4: Nominative subject + non-inflected infinitive

*O pai deixou / viu elas beber sumo.
 the father let / saw they.NOM drink.INF juice
 “The father let/saw them drink juice.”

Regarding the second AJT, it focuses on PIC and tests the acquisition of Case and ϕ -features of the embedded domain in each structure as well. We also defined four conditions: (i) Accusative subject + prepositional inflected infinitive (PIC); (ii) Accusative subject + prepositional uninflected infinitive (PIC); (iii) Nominative subject + prepositional inflected infinitive (ungrammatical); (iv) Nominative subject + prepositional uninflected infinitive (ungrammatical). Once again, we had 16 test items (4 per condition, with the perception verb *ver* “to see”). Table 2 summarizes the conditions tested.

Table 2. Conditions of the AJT testing PIC

Condition 1: Accusative subject + inflected prepositional infinitive (PIC)

^{OK}A mãe viu-os a limparem a cozinha.
 the mother saw-CL.ACC to.ASP clean.INF.3PL the kitchen
 “The mother saw them cleaning the kitchen.”

Condition 2: Accusative subject + non-inflected prepositional infinitive (PIC)

^{OK}A mãe viu-as a pentear os cabelos.
 the mother saw-CL.ACC to.ASP comb.INF the hair
 “The mother saw them combing their hair.”

Condition 3: Nominative subject + inflected prepositional infinitive

*O professor viu elas a arrumarem os livros.
 the teacher saw they.NOM to.ASP pack.INF.3PL the books
 “The teacher saw them packing the books.”

Condition 4: Nominative subject + non-inflected prepositional infinitive

*A Carla viu eles a comprar os bilhetes.
 the Carla saw they.NOM to.ASP buy.INF the tickets
 “Carla saw them buying the tickets.”

Both AJT presented a 2:1 ratio of distractor sentences, following Sprouse (2007) and Kraš (2010). We also controlled the properties of the embedded subjects. They are all 3rd person plural so that there is overt subject-verb agreement morphology (in the conditions with an inflected infinitival verb form), pronominal, to make visible their Case marking, and [+human], in order to identify transfer effects related to DOM.

Participants were asked to evaluate each sentence using a four-point Likert scale (0 to 3), plus the “I don’t know” option. They were also asked to correct the sentences they considered ungrammatical in EP. The task was completed in two sessions, presented in paper format, with no time limit and was preceded by a text explaining the task. The participants also filled out a form with information about their sociolinguistic profile.

The data from the test group were collected at the University of Extremadura and *Escuelas Oficiales de Idiomas* (Official Schools of Languages) in Spain, in the autonomous communities of Castilla y León and Extremadura, and in Madrid. All participants were learning EP in a formal learning context. Taking into account the dialectal variation that characterizes the Spanish territory, we restricted our data collection to those geographical areas and excluded data from every bilingual participant of other official or minority languages spoken in Spain. We also excluded any participants who started learning EP during childhood. We collected data from Spanish learners of EP with different (completed) proficiency levels at the time of data collection: A2 – initial level ($N = 22$); B1 – intermediate level ($N = 26$); B2 – advanced level ($N = 11$). We also included a control group of native (monolingual) EP speakers with no background in Linguistics ($N = 43$). The participants from both groups were between 18 and 49 years of age. The two groups contained more female than male participants. In the case of the non-native EP learners, 69.5% were female and 30.5% were male. As for the monolingual EP speakers, 79% were female and 21% were male.

5. Acceptability judgment tasks: Data analysis

In this section, we report the results obtained from the two AJT described before. We conducted a statistical analysis of the data collected using SPSS (version 25). We used non-parametric statistic tests, more specifically, a Kruskal-Wallis test with Bonferroni correction for multiple comparisons to analyse the answers of the participants entered as an ordinal variable (0 to 3) in each condition. The cases corresponding to the choice of “I don’t know” were removed from the analysis.

5.1 Exceptional case marking and inflected infinitive structure

The results from the first AJT, i.e., the task concerning ECM and IIS, show a clear difference in the judgments provided by L2 EP learners and EP native speakers for the conditions with Accusative subjects and the ones with Nominative subjects.

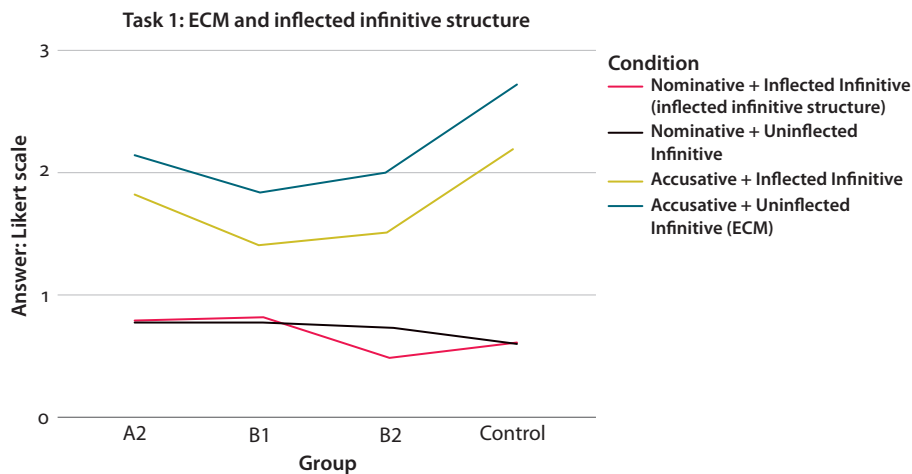


Figure 1. Results from the AJT testing ECM and IIS

Figure 1 shows a strong tendency for rejecting sentences with pronominal Nominative subjects across all groups. In fact, the statistical analysis revealed that there are no significant differences between groups in the conditions corresponding to IIS, i.e., Nominative + inflected infinitive ($H(3) = 7.525, p = .057$) and to Nominative subject + uninflected infinitive ($H(3) = 5.158, p = .161$).

We indeed predicted that Spanish learners of L2 EP would struggle with the reassembly of the feature bundles of IIS, especially in less advanced levels of proficiency, which would result in their rejection of this structure. However, the interpretation of the data should be revised in light of the results of native speakers, who also rejected the IIS in this context, even though it is described as part of the EP standard grammar. Similar data for IIS (with pronominal subjects) have been reported by Barbosa et al. (2018), as previously mentioned.

To understand the rejection of the IIS structure, we must consider the type of corrections provided by native speakers, since the majority of the corrections given for rejected IIS sentences correspond to ECM with inflected infinitive (42%) or, in the case of perception verbs, PIC (with or without inflected infinitive) (24%). On the other hand, L2 learners offer a wider range of types of corrections, either selecting structures available in their L1 (e.g., ECM, *faire-inf*, finite complement

clauses), or changing the Nominative for an Accusative or Dative pronoun. We discuss this preference for a Dative pronoun below. Previous studies on L1 acquisition of infinitival complement clauses, including in the contexts under study, show that monolingual EP speakers (both children and adults) produce inflected infinitives (mainly with full DP subjects). This, along with our data, seems to indicate that the problem does not lie in the inflected infinitive verbal form itself, but in the Case of the embedded subject.

On the other hand, the sentences with Accusative pronominal subjects achieved higher acceptance rates, with all groups showing a similar tendency in this regard. Despite this general tendency, we found significant differences between groups in ECM ($H(3) = 48.258, p < .001$) and in the Accusative + inflected infinitive ($H(3) = 39.160, p < .001$) conditions. The pairwise comparisons conducted revealed differences between the control group and all the L2 EP learner groups in these conditions: ECM (all $p < .05$) and Accusative + inflected infinitive (all $p < .05$). In fact, if we look closely at Figure 1, we can see that whereas native speakers clearly accept the Accusative + inflected infinitive structure, the L2 EP learners show less clear results, with B1 and B2 speakers being closer to rejection. The acceptance of this structure (with 3rd person plural subjects) by native EP speakers is in line with the results previously reported by Barbosa et al. (2018). Hence, our data seems to add evidence that supports the hypothesis of this structure being available in EP.

Considering now as a whole the acceptability rates of the IIS and the Accusative + inflected infinitive in the case of judgments provided by EP native speakers, we hypothesised that a task effect cannot be excluded: the present task is an Acceptability Judgment Task, in which hyper-correction effects can be found. To test this hypothesis, we gathered data from corpora, trying to attest both structures. We collected data from two different corpora: one written (CETEMPúblico) and one including both oral and written texts (CRPC – Reference Corpus of Contemporary Portuguese).¹³ After a search with causative and perception verbs, we found a very low frequency of the IIS (only 13 occurrences in CETEMPúblico and 8 occurrences in CRPC), but similarly low frequency of the Accusative + inflected infinitive structure (12 occurrences in CETEMPúblico and 9 in CRPC). Only cases in which the relevant DPs were pronominal, and therefore presented overt Case, were considered. Since these numbers do not reflect the differences in the acceptability of the two structures that we found, we cannot exclude a task effect.

13. CETEMPúblico is a corpus containing about 180 million words in EP, available at <https://www.linguateca-pt/CETEMPUBLICO/>. CRPC contains over 289 million words of European Portuguese texts and is available through <https://clul.ulisboa.pt/en/recurso/reference-corpus-contemporary-portuguese>.

Furthermore, it comes as no surprise that ECM (Accusative + uninflected infinitive) is the structure with higher acceptance rates. However, at first sight, the fact that significant differences between native speakers and all the L2 learner groups were found (even in advanced levels of proficiency) is puzzling. If we consider that Spanish learners of L2 EP would only need to map the properties of a defective T to acquire ECM in EP, we would expect no such differences to be found, at least in the B2 group. An explanation for these data can be found in the corrections provided for the rejected sentences: the majority (close to 75%) of the corrections consists in the replacement of the Accusative subject by a Dative. This may be caused by lingering transfer effects of the L1 properties regarding the Case marking of constituents in Object position. In other words, we argue that Spanish speakers tend to grammaticalize (Longobardi 2005) the semantic features associated with DOM even if this is not possible in the L2. This seems to be in line with the assumption that avoiding features not encoded in the L2 is a complex task.

5.2 Prepositional infinitival construction

Regarding the AJT focused on PIC, the results are similar to the ones reported for the ECM and IIS, in that there seems to be a dividing line between the acceptance rates of sentences with Accusative and with Nominative subjects, with the latter obtaining worse results, even though this difference is more expressive in the case of the control group. In the case of this task, a general rejection of nominative was expected in the control group, as shown in Figure 2.

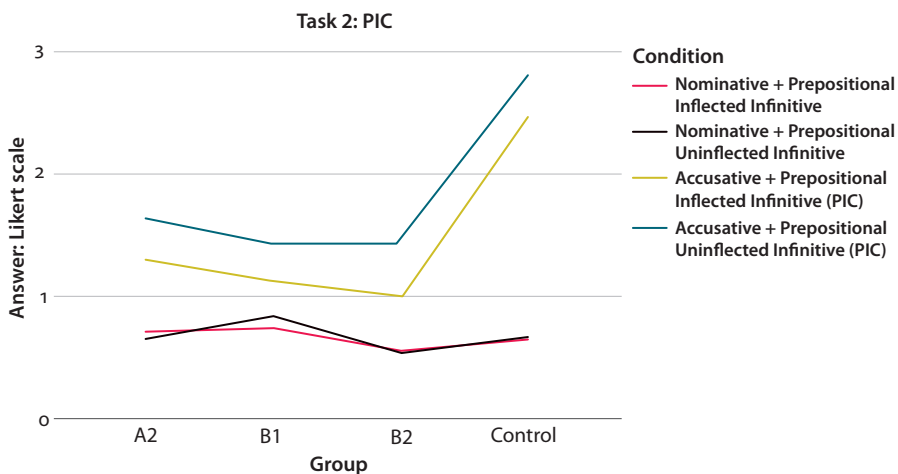


Figure 2. Results from the AJT testing PIC

There is a clear rejection of the conditions with Nominative subjects across all groups, which is expected. The statistical analysis confirms this, for there are no differences between groups either in the condition with Nominative + inflected infinitive ($H(3) = 2.536, p < .469$) or with Nominative + uninflected infinitive ($H(3) = 6.463, p < .091$).

As for the conditions with Accusative subjects, i.e., the conditions corresponding to PIC with inflected and uninflected infinitive, the results for the control group are clear: native EP speakers accept both instances of PIC. This result is also in line with Santos, Gonçalves & Hyams (2016), whose study has shown that PIC seems to be the complement of perception verbs preferred by EP monolingual speakers (children and adults). However, the results obtained for the control group contrast with those of L2 EP learners. The latter show lower acceptance rates of PIC (on the verge of rejection, especially in the case of PIC with inflected infinitive). The statistical test reflects this since there are significant differences between groups for the PIC with inflected infinitive ($H(3) = 129.603, p < .001$) as well as for the PIC with uninflected infinitive ($H(3) = 139.290, p < .001$). Pairwise comparisons showed that there are significant differences between the control group and all the L2 EP learner groups for PIC with inflected infinitive (all $p < .001$) and uninflected infinitive (all $p < .001$).

The corrections provided by the Spanish L2 EP learners shed some light on the reasons why PIC gets low acceptance rates. Crucially, Spanish speakers do not replace the rejected instances of PIC for GC. This type of correction was found in a residual number of occurrences, all from the same participant. Instead, a wide range of corrections is provided, from which we highlight the more significant. First, we registered cases of replacement of PIC for instances of ECM (29%) or ECM with inflected infinitive (13%). In such cases, the rejection of the PIC seems to be motivated by the presence of the aspectual marker *a* “to”. On the other hand, some corrections can be related, once again, to transfer effects of the DOM properties from the L1, since we found a number of replacements of Accusative for Dative subjects (36%). Here we found instances of Dative subjects co-occurring with bare infinitives, prepositional infinitives, and inflected infinitives.

The absence of corrections involving GC seems to point to a difficulty in mapping the aspectual progressive value associated with the Gerund verb form in the L1 to the aspectual head *a* “to” plus infinitive verb form of the PIC, or, assuming Casalicchio’s (2019) analysis, Spanish learners of L2 EP seem to struggle with mapping the incorporated aspectual marker in the Gerund verb form in the L1 with the spelled-out aspectual marker of the PIC in the L2. This result agrees with our initial predictions: we predicted that the difficulty in acquiring the PIC in the case of Spanish native speakers starts in the mapping stage and that it could be explained by the difficulty in identifying a formal contrast associated with the presence or absence of the preposition *a* “to” preceding an infinitival verb form.

6. Conclusions

Overall, our data seems to be in line with the main predictions of the FRH and with the particular predictions we formulated based on the FRH. As shown in previous sections, L2 EP learners struggle with the feature bundles that need to undergo reassembly in the L2.

In the case under study, and considering the differences and similarities between Spanish and EP, whereas Spanish speakers seem to acquire the properties that only require mapping earlier, they seem to struggle, even in advanced levels of proficiency, with the reassembly of features that are not available or that are bundled in different configurations in the L1 and in the L2. The results reported for ECM seem to support this, since Spanish learners of EP seem to be able to map the features associated with the defective T in ECM, but show difficulty in dissociating (or not grammaticalizing) the semantic features related to DOM in the Case marking of the embedded subject.

As for the IIS, the results obtained for the monolingual EP speakers point to the necessity of gathering more data to understand the rejection of the structure in this particular task. On the other hand, the positive results obtained for the structure with Accusative subject and inflected infinitive (or ECM with inflected infinitive) are a clear indication that this structure must be included in the discussion about L2 acquisition of infinitival complements of causative and perception verbs. Once again, if we take the acceptance rates for this structure (obtained for Spanish L2 EP learners) together with the ones reported for PIC with inflected infinitive, we see that Spanish speakers struggle with different instances of inflected infinitive, in contrast to clear acceptance of structures involving inflected infinitives in the native EP data.

This is particularly evident in the PIC data. In fact, Spanish learners show very low acceptance rates of PIC with inflected and uninflected infinitive. These results, along with the fact that participants did not correct the rejected instances of PIC for GC, confirm our prediction that the locus of difficulty in the acquisition of PIC for Spanish native speakers lies in the mapping of the aspectual progressive value, since the aspectual features of the PIC and their representation in the L2 are not perceived as close to the ones of the GC, in the L1.

Acknowledgements

We would like to thank the participants in the experiment, as well as the schools and teachers who accepted to participate in this study.

Funding

This research was funded by Fundação para a Ciência e a Tecnologia (grant FCT/SFRH/BD/92581/2013 and funding attributed to the Center of Linguistics of the University of Lisbon, UIDB/00214/2020).

References

- Ambar, Manuela. 1992. *Para uma sintaxe da inversão sujeito-verbo em Português*. Lisbon: Edições Colibri.
- Barbosa, Pilar & Fernanda Cochofel. 2005. A construção de infinitivo preposicionado em PE. In Inês Duarte & Isabel Leiria (eds.), *Actas do XX Encontro Nacional da APL*, 387–400. Lisbon: APL/Edições Colibri.
- Barbosa, Pilar, Cristina Flores & Cátia Pereira. 2018. On subject realization in infinitival complements of causative and perceptual verbs in European Portuguese. Evidence from monolingual and bilingual speakers. In Pedro Guijarro-Fuentes & Alejandro Cuza (eds.), *Language acquisition and contact in the Iberian Peninsula*, 125–158. Boston: Mouton de Gruyter. <https://doi.org/10.1515/978151059988-006>
- Barbosa, Pilar & Eduardo P. Raposo. 2013. Subordinação argumental infinitiva. In Eduardo P. Raposo, Maria Fernanda Bacelar do Nascimento, Maria Antónia Coelho Mota, Luísa Segura & Amália Mendes (eds.), *Gramática do Português*, Volume 2, 1901–1977. Lisbon: Fundação Calouste Gulbenkian.
- Bordelois, Ivonne. 1988. Causatives: From lexicon to syntax. *Natural Language and Linguistic Theory* 6, 57–93. <https://doi.org/10.1007/BF01791592>
- Brugè, Laura & Gerhard Brugger. 1996. On the accusative *a* in Spanish. *Probus* 8, 1–51. <https://doi.org/10.1515/prbs.1996.8.1.1>
- Casalichchio, Jan. 2013. Pseudorelative, gerundi e infiniti nelle varietà romanze: Affinità solo superficiali e corrispondenze strutturali. Ph.D. Dissertation: Università degli Studi di Padova.
- Casalichchio, Jan. 2019. Gerunds become prepositional infinitives in Romance Small Clauses: The effects of later Merge to the syntactic spine. *Probus* 31(1), 75–117. <https://doi.org/10.1515/probus-2019-0003>
- Castillo, Concha. 2001. The configuration of ECM structures. *Studia Linguistica* 55(2), 113–140. <https://doi.org/10.1111/1467-9582.00076>
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos P. Otero & Maria-Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory*, 291–321. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9780262062787.003.0007>
- Contreras, Heles. 1987. Small clauses in Spanish and English. *Natural Language and Linguistic Theory* 5, 225–245. <https://doi.org/10.1007/BF00166585>
- Demirdache, Hamida & Myriam Uribe-Etxebarria. 1997. The syntax of temporal relations: A uniform approach to tense and aspect. *Proceedings of the Sixteenth West Coast Conference on Formal Linguistics, WCFL 16*, 145–159. Stanford, CA: CSLI Publications.
- Duarte, Inês. 1992. Complementos infinitivos preposicionados e outras construções temporalmente defectivas em Português Europeu. *Actas do VIII Encontro da APL*, 145–158. Lisbon: APL/Edições Colibri.

- Duarte, Inês. 2003. Subordinação completiva – as orações completivas. In Maria Helena M. Mateus, Ana Maria Brito, Inês Duarte, Isabel H. Faria, Sónia Frota, Gabriela Matos, Fernanda Oliveira, Marina Vigário & Alina Villalva (eds.), *Gramática da língua Portuguesa*, 5th edn., 593–651. Lisboa: Caminho.
- Duarte, Inês & Anabela Gonçalves. 2002. Construções de subordinação funcionalmente defecivas: o caso das construções perceptivas em Português Europeu e em Português Brasileiro. In Anabela Gonçalves & Clara N. Correia (eds.), *Actas do XVII Encontro da APL*, 161–173. Lisbon: APL/Edições Colibri.
- Duarte, Inês, Anabela Gonçalves & Matilde Miguel. 2005. Propriedades de C em frases completivas. In Inês Duarte & Isabel Leiria (eds.), *Actas do XX Encontro Nacional da APL*, 549–562. Lisbon: APL.
- Gonçalves, Anabela. 1999. Predicados complexos verbais em contextos de infinitivo não preposicionado do português europeu. Ph.D. Dissertation: University of Lisbon.
- Gonçalves, Anabela & Inês Duarte. 2001. Construções causativas em Português Europeu e em Português do Brasil. In Clara N. Correia & Anabela Gonçalves (eds.), *Actas do XVI Encontro da APL*, 657–671. Lisbon: APL.
- Gonçalves, Anabela, Ana Lúcia Santos & Inês Duarte. 2014. (Pseudo-) Inflected infinitives and Control as Agree. In Karen Lahousse & Stefania Marzo (eds.), *Selected papers from 'Going Romance' Leuven 2012*, 161–180. Amsterdam: John Benjamins.
<https://doi.org/10.1075/rllt.6.08gon>
- Hale, Ken. 1986. Notes on world view and semantic categories: Some Walpiri examples. In Pieter Muysken & Henk van Riemsdijk (orgs.), *Features and projections*, 233–254. Dordrecht: Foris. <https://doi.org/10.1515/9783110871661-009>
- Hawkins, Roger & Hajime Hattori. 2006. Interpretation of English multiple wh-questions by Japanese speakers: A missing uninterpretable feature account. *Second Language Research* 22. 269–301. <https://doi.org/10.1191/0267658306sr2690a>
- Hernanz Carbó, María L. 1999. El infinitivo. In Ignacio Bosque & Violeta Demonte (eds.), *Gramática descriptiva de la Lengua Española*, Volume 2. Madrid: Real Academia Española/Espasa Calpe.
- Hornstein, Norbert, Ana Maria Martins & Jairo Nunes. 2008. Perception and causative structures in English and European Portuguese: ϕ -feature agreement and the distribution of bare and prepositional infinitives. *Syntax* 11(2). 205–229.
<https://doi.org/10.1111/j.1467-9612.2008.00105.x>
- Kraš, Tihana. 2010. Unaccusativity in L2 Italian at the lexicon-syntax interface. In Michael Iverson, Ivan Ivanov, Tiffany Judy, Jason Rothman, Roumyana Slabakova & Marta Tryzna (eds.), *Proceedings of the 2009 Mind/Context Divide Workshop*, 60–71. Somerville, MA: Cascadilla Proceedings Project.
- Lardiere, Donna. 2000. Mapping features to forms in second language acquisition. In John Archibald (ed.), *Second language acquisition and linguistic theory*, 102–129. Oxford: Blackwell.
- Lardiere, Donna. 2008. Feature-assembly in second language acquisition. In Juana Liceras, Helmut Zobl & Helen Goodluck (eds.), *The role of formal features in second language acquisition*, 106–140. New York: Lawrence Erlbaum.
- Lardiere, Donna. 2009. Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research* 25. 173–227.
<https://doi.org/10.1177/0267658308100283>
- Longobardi, Giuseppe. 2005. A minimalist program for parametric linguistics. In Hans Broekhuis, Norbert Corver, Riny Huybregts, Ursula Kleinhenz, Jan Koster (eds.), *Organizing grammar: Linguistic studies in honor of Henk van Riemsdijk*, 407–414. Berlin: Mouton de Gruyter.

- Madeira, Ana Maria. 1994. On the Portuguese inflected infinitive. *UCL Working Papers in Linguistics* 6. 180–203.
- Martins, Ana Maria. 2018. Infinitival complements of causative/perception verbs in a diachronic perspective. In Ana Lúcia Santos & Anabela Gonçalves (eds.), *Complement clauses in portuguese. Syntax and acquisition*, 101–128. Amsterdam: John Benjamins.
<https://doi.org/10.1075/iuhl.17.04mar>
- Mensching, Guido. 2000. *Infinitive constructions with specified subjects. A syntactic analysis of Romance Languages*. Oxford: Oxford University Press.
- Montrul, Silvina & James Yoon. 2009. Putting parameters in their proper place. *Second Language Research* 25(2). 291–311. <https://doi.org/10.1177/0267658308100290>
- Pereira, Sílvia A. 2015. Predicative constructions with gerunds in European Portuguese dialects. *Dialectologia*. Special Issue 5. 351–371.
- Rafel, Joan. 2000. Complex small clauses. Ph.D. Dissertation: Universitat Autònoma de Barcelona.
- Raposo, Eduardo P. 1981. A construção “União de orações” na gramática do Português. Ph.D. Dissertation: University of Lisbon.
- Raposo, Eduardo P. 1987. Romance infinitival clauses and case theory. In Carol Neidle & Rafael A. Nuñez Cedeño (eds.), *Studies in Romance Languages*, 237–250. Dordrecht: Foris.
<https://doi.org/10.1515/9783110846300.237>
- Raposo, Eduardo P. 1989. Prepositional infinitival constructions in European Portuguese. In Osvaldo Jaegli & Kenneth J. Safir (eds.), *The null subject parameter*, 277–305. Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-009-2540-3_10
- Rodríguez-Mondoñedo, Miguel. 2008. The acquisition of differential object marking in Spanish. *Probus* 20. 1–31. <https://doi.org/10.1515/PROBUS.2008.004>
- Rothman, Jason & Roumyana Slabakova. 2018. The generative approach to SLA and its place in modern second language studies. *Studies in Second Language Acquisition* 40(2). 417–442.
<https://doi.org/10.1017/S0272263117000134>
- Santos, Ana Lúcia, Anabela Gonçalves, & Nina Hyams. 2016. Aspects of the acquisition of object control and ECM-type verbs in European Portuguese. *Language Acquisition* 23(3). 199–233.
<https://doi.org/10.1080/10489223.2015.1067320>
- Sprouse, Jon. 2007. A program for experimental syntax: Finding the relationship between acceptability and grammatical knowledge. Ph.D. Dissertation: University of Maryland.
- Torrego, Esther. 1998. *The dependencies of objects*. Cambridge, MA: MIT Press.
<https://doi.org/10.7551/mitpress/2337.001.0001>
- Torrego, Esther. 2010. Variability in the case patterns of causative formation in Romance and Its Implications. *Linguistic Inquiry* 41(3). 445–470. https://doi.org/10.1162/LING_a_00004
- Treño, Estela. 1994. *Las causativas del español con complemento de infinitivo*. Mexico City: El Colegio de Mexico.
- Tsimplí, Ianthi M. & Maria Dimitrakopoulou. 2007. The interpretability hypothesis: Evidence from wh-interrogatives in second language acquisition. *Second Language Research* 23(2). 215–242. <https://doi.org/10.1177/0267658307076546>
- Tubino, Mercedes. 2010. Contrasting causatives: A minimalist approach. Ph.D. Dissertation: University of Arizona.

Focus fronting vs. *wh*-movement

Evidence from Sardinian

Silvio Cruschina

University of Helsinki

It is commonly held that focus fronting exhibits similar properties to *wh*-movement. The syntactic parallelism between the two types of movement has been supported by the semantic analyses of *wh*-questions that assume that when *wh*-words function as interrogative operators, they are inherently focal. The main goal of this paper is to challenge this highly attractive picture of the relationship between *wh*-words, focus, and movement, and to claim that *wh*-phrases are not inherently focal. The results of a prosodic production experiment on the distribution of the nuclear pitch accent in Sardinian *wh*-questions, together with the syntactic properties related to the asymmetry between direct and indirect *wh*-questions, form the empirical basis of this study.

Keywords: focus fronting, *wh*-question, nuclear pitch accent, Sardinian, production experiment

1. Introduction

It is a commonly held view that focus fronting (henceforth FF) exhibits similar properties to *wh*-movement. A comparison of these two operations, however, reveals clear syntactic differences. In English, for example, *wh*-words undergo (obligatory) movement in *wh*-questions, but FF is not obligatory and is instead limited to special interpretations. Relying on evidence from weak crossover effects, however, Chomsky (1976) proposes a covert (LF) movement operation for FF. Even if the focus stays in situ (1b), the configuration gives rise to the same type of weak crossover effects as observed in *wh*-questions (2):

- (1) a. His_i mother saw $John_i$
b. * His_i mother saw JOHN_i
- (2) *Who_i did his_i mother see?

Sentences with narrow focus thus constitute an operator-variable structure, meaning that the representation of focus at LF is parallel to that of structures derived by *wh*-movement. The observable syntactic difference between *wh*-movement and FF can therefore be expressed simply in terms of a difference in the nature of movement: overt vs. covert movement. Consequently, several treatments of question syntax assume a parallelism between *wh*-phrases and foci, and suggest that *wh*-phrases carry (or at least can carry) a focus feature (see, e.g., Haegeman & Guéron 1999).

According to a longstanding tradition starting with Brody (1995) and Rizzi (1997), it has additionally been assumed that *wh*-movement and FF target the same position within the left periphery of the clause. This position is the specifier of FocP in Rizzi (1997), as shown in (3):

- (3) [Force [Top* [Foc [Top* [Fin [IP...]]]]]]

The syntactic parallelism between the two types of movement has been further supported by the semantic analyses of *wh*-questions that assume that when *wh*-words function as interrogative operators, they are inherently focal (Beck 2006; Cable 2010; see also Eckardt 2007 for discussion). Beck (2006), for example, assumes that interrogative *wh*-phrases only contribute a focus semantic value to the interpretation (4), which is exploited by a question operator on top of the sentence radical:

- (4) [What]_F *did Dave buy?*
 Focus Background/Presupposition

The main aim of this paper is to challenge this highly attractive picture of the relationship between *wh*-words, focus, and movement. My starting point is Jones's (1993) comparison between fronted foci and *wh*-phrases in Sardinian and, in particular, the prosodic differences between *wh*-questions and sentences with FF, as well as the more general asymmetry between direct and indirect *wh*-questions. Following Bianchi, Bocci & Cruschina (2017), I propose that interrogative *wh*-words are not inherently focal, and that direct (but not indirect) *wh*-questions require a congruent answer, and hence disallow an independent focus structure.

In § 2, I briefly review Jones's (1993) discussion of the differences and similarities between FF and *wh*-movement in Sardinian. I then concentrate on the prosodic properties of these two structures, commenting first on the general problems that the distribution of the 'nuclear pitch accent' (NPA) in *wh*-questions poses for any theory that assumes a parallelism between *wh*-phrase and focus (§ 3), and then presenting the results of a prosodic experiment on Sardinian *wh*-questions (§ 4). In § 5, I discuss a further important difference between *wh*-questions and sentences with FF, namely, the difference between direct and indirect *wh*-questions in terms of syntactic restrictions: a *wh*-phrase can only co-occur with a fronted focus in

indirect questions, and is ruled out in direct questions. In § 6, I outline an analysis that aims to account for both the root-embedded asymmetry with respect to *wh*-questions and the distribution of the NPA in Sardinian. This analysis is based on the claim that *wh*-questions are not inherently focal. The main points of the paper are summarized in § 7.

2. FF vs. *wh*-movement in Sardinian

Sardinian features an extensive use of FF, which is not limited to a contrastive interpretation as in other Romance languages. This allows a more direct comparison between the two movement types: FF and *wh*-movement. Jones (1993: 333–334) lists a number of common properties between the two structures: (a) uniqueness, i.e., movement can only apply to one element within the same clause; (b) adjacency between the fronted item and the verb, which prevents the subject from occurring in a preverbal position; (c) ordering with respect to left-dislocated phrases; (d) incompatibility with the interrogative particle *a*; and (e) incompatibility with each other, in that fronted foci and *wh*-phrases are mutually exclusive – at least in root clauses (see also Jones 2013). The last of these is illustrated in (6), where it is shown that *wh*-movement and FF, which can occur separately (cf. (5a) and (5b), respectively), cannot occur within the same sentence, irrespective of the linear order (6):

- (5) a. *A kie as datu su jocátulu?*
 to whom have.PRS.2SG given the toy
 “To whom did you give the toy?”
 b. *Su jocátulu li as datu.*
 the toy him.DAT= have.PRS.2SG given
 “You gave him *the* toy.”
- (6) a. **A kie su jocátulu as datu?*
 to whom the toy have.PRS.2SG given
 b. **Su jocátulu a kie as datu?*
 the toy to whom have.PRS.2sg given

Jones accounts for the mutual incompatibility between fronted foci and *wh*-items, as well as for other properties, by postulating that both types of movement target the same functional projection in the left periphery of the clause – as later developed in Rizzi (1997).

However, there are clear differences between the two types of movement: (a) position of the stress; (b) occurrence in embedded contexts, which applies to FF, but only applies to *wh*-phrases in indirect questions; (c) availability in embedded

infinitival clauses, which applies only to indirect *wh*-questions, and not to FF; (d) presence of a preverbal subject position in embedded clauses, which is only an option in indirect *wh*-questions.

After reviewing the similarities between FF and *wh*-movement in Sardinian, Jones (1993) begins his discussion of the differences with an observation about the different position of the stress in the two types of structure:

On the other hand, there are a number of differences between NP fronting and WH-movement. First, whereas fronted NPs are always stressed and focused, to the exclusion of other elements in the remainder of the sentence, *wh*-items do not normally bear the main stress except in echo-questions or in cases where the content of the rest of the clause is pragmatically presupposed. (Jones 1993: 335)

Unlike fronted foci, thus, *wh*-phrases do not bear the NPA of the sentence (the main stress, in Jones's terminology). Moreover, fronted foci and *wh*-phrases display different behaviour in embedded contexts: as in Italian (Rizzi 2001; Bocci, Rizzi & Saito 2018) and in other Romance languages, *wh*-phrases are incompatible with a narrow focus in direct questions, but not in indirect questions (7):

- (7) *Mi so dimandande a Juanne itte appan natu* (,no a Paulu).
 me= be.1SG ask.GER to Juanne what had.3PL said not to Paulu
 "I'm wondering what they said to Juanne (, not to Paulu)."

Unlike in direct questions, then, in embedded (indirect) questions FF and *wh*-movement are compatible with each other. In this paper I primarily concentrate on the differences in prosodic patterns between *wh*-questions and sentences with FF, and will then further comment on the asymmetry between direct and indirect *wh*-questions with respect to the mutual (in)compatibility between fronted foci and *wh*-phrases. For the purposes of the first of these, in § 4 I present the results of a prosodic experiment confirming Jones's intuitions about the different prosodic pattern of *wh*-questions, and undermining the received 'isomorphism' between focus fronting and *wh*-movement.

3. The distribution of the 'nuclear pitch accent' in *wh*-questions

According to a well-established approach, focus is always realized by structural prominence. This approach was pioneered by Truckenbrodt (1995) with the constraint called Focus Prominence given in (8):

- (8) Focus Prominence (Truckenbrodt 1995)
Focus needs to be maximally prominent

In order to be maximally prominent, the focus constituent of the sentence must be assigned the most prominent metrical head of the sentence, i.e., the nuclear stress (Zubizarreta 1998, 2016; Zubizarreta & Vergnaud 2005); at the intonational level, this metrical head is in turn associated with the pitch accent that qualifies as the NPA (Beckman, Hirschberg & Shattuck-Hufnagel 2005).

In Romance, if no narrow focus is present, the NPA is assigned by default to the rightmost constituent of the sentence in declaratives. In *wh*-questions, however, a number of different patterns are found.

3.1 The prosodic patterns of *wh*-questions

Ladd (1996: 170–174) observes the curious exceptional behaviour of *wh*-questions with respect to the assignment of the NPA: “Logic seems to suggest that the WH-word is the focus of the question, and yet, in English at least, the WH-word does not normally bear the most prominent accent” (the NPA, in our terminology). English displays the pattern in (9a), with the NPA on the stressed syllable of the rightmost constituent, rather than the expected contour in (9b) with the NPA on the *wh*-phrase:¹

- (9) a. *Where are you **going**?*
 b. *#Where are you going?*

From a comparative viewpoint, Ladd (1996: 171) remarks that “there are languages with WH-movement that do put the nuclear accent on the WH-word”. One such language is Romanian, for which he provides the following examples (the same pattern is also found in Hungarian and Greek):

- (10) (Romanian)
 a. ***Unde** mergi?* “Where are you going?”
 b. ***Când** a plecat?* “When did he leave?”
 c. ***Cine** a chemat?* “Who called?”

According to Ladd, there are therefore two basic typological patterns:

- i. *wh*-questions follow the same stress principles as statements, that is, the NPA is on the stressed syllable of the rightmost constituent (as in English);
- ii. *wh*-questions are treated differently from statements, in that the NPA is on the *wh*-word (as in Romanian, Hungarian, and Greek).

1. In the examples from (9) to (11), I use bold to indicate the syllable associated with the NPA. In the original examples, Ladd (1996: 170–174) employs capital letters for the same purpose.

Ladd further observes that there is a difference between Romanian and Italian in the assignment of the NPA in *wh*-questions: while the NPA is on the *wh*-word in Romanian (11a), it falls on the verb in Italian (11b):

- (11) H* L L%
- a. *Unde l'ai cumpărat?* (Romanian)
 where it=has.2SG bought
 “Where did you buy it?”
 H* H+L* LL%
- b. *Dove l'hai comprato?* (Italian)
 where it=has.2SG bought
 “Where did you buy it?”

It thus seems that Italian *wh*-questions behave differently.² Compelling evidence for the distinct prosodic behaviour of Italian *wh*-questions comes from Marotta (2001) and Bocci, Bianchi & Cruschina (2021), who, with the aid of experimental evidence, show that the NPA is assigned to the lexical verb in Italian:³

- (12) *Chi canta una canzone?*
 who sing.PRS.3SG a song
 “Who is singing a song?” (Marotta 2001)
- (13) *Chi pensa che ti dovrei presentare al direttore?*
 who think.PRS.3SG that you= should.1SG introduce.INF to-the director
 “Who thinks that I should introduce you to the director?”
 (Bocci et al. 2021: 417 Example 19a, 421 Figure 2)

Bocci et al. (2021) additionally point out that the systematic assignment of the NPA to the lexical verb observed in Italian direct *wh*-questions does not seem to emerge

2. In fact, since the verb is the last constituent of the sentence, the example in (11b) could be taken to illustrate the same pattern as English, where the NPA is on the stressed syllable of the rightmost constituent. Ladd's observations about Italian are therefore not entirely clear. At any rate, in the later studies by Marotta (2001) and Bocci et al. (2021) the verb is followed by another constituent (e.g., a direct object or a complement clause), hence ensuring that the accented verb does not occupy the rightmost position, and more clearly showing that the Italian pattern is indeed a distinct type.

3. This generalization about the placement of the NPA in Italian only holds for *wh*-questions featuring a bare *wh*-element (i.e., the equivalents of *who*, *what*, *where*, and *how*). The interrogative element *perché* “why” constitutes an exception to this pattern because it can be assigned the NPA, as shown by Marotta (2001). The distinct behaviour of *perché* compared to other *wh*-phrases is not completely unexpected given that it also displays different syntactic properties (see Rizzi 2001; Bianchi et al. 2017; Bocci & Cruschina 2018). Complex *wh*-phrases, such as D-linked *wh*-phrases, also seem to behave differently, in that they tend to attract the NPA in Italian (see Bocci, Cruschina & Rizzi 2021, for some preliminary experimental data).

in other clause types, and that this consistent placement does not seem to be related to focus from an interpretive viewpoint: indeed, the lexical verb is not interpreted as focussed in (12) or in (13).⁴ If we add the Italian pattern to Ladd’s typology, we have three patterns, which I label WH, V, and R for convenience (V and R stand for V(erb) and R(ightmost), respectively):

- i. WH: the NPA is on the stressed syllable of the *wh*-phrase;
- ii. V: the NPA is on the stressed syllable of verb;
- iii. R: the NPA is on the stressed syllable of the rightmost constituent.

Interestingly, all three patterns are attested in Romance. We know already that Romanian displays the WH pattern, while Italian exhibits the V pattern. By contrast, the R pattern is found in Western Romance, for example in Spanish, Catalan, and in Portuguese (Hualde & Prieto 2015; Prieto et al. 2015; Frota et al. 2015). An example from Spanish is given in (14):

- (14) L+<H* H* L* H%
- ¿Dónde está tu gasolinera?
- where is your gas station
- “Where is your gas station?” (Hualde & Prieto 2015: 382)

So, what pattern do we find in Sardinian? This question will be addressed in § 4 with evidence from a prosodic experiment. Let us first review the literature on the prosody of Sardinian *wh*-questions.

3.2 The NPA in Sardinian *wh*-questions

As mentioned in § 2, Jones (1993: 355) acknowledges the difference between the prosodic behaviour of *wh*-questions and that of FF: “whereas fronted NPs are always stressed and focused, ..., *wh*-items do not normally bear the main stress except in echo-questions or in cases where the content of the rest of the clause is pragmatically presupposed”.⁵ An example of the contrast is discussed in his later work and is reproduced here as (15):

4. The results of the prosodic experiment reported in Bocci et al. (2021) show that the NPA nearly always falls on a lexical verb. The NPA is never (0%) assigned to the rightmost element of the sentence, and the association of the NPA with the *wh*-phrase is very marginal (less than 2%). See Bocci et al. (2021) for more details.

5. Although I will not address the special behaviour of echo questions in this paper, it must be noted that, cross-linguistically, the focal status of the *wh*-element in echo questions is less controversial, both prosodically and syntactically. By contrast, however, echo *wh*-questions are not proper interrogative sentences and might not even involve a [wh] feature at all (see Reis 1991, 2012; see also Sobin 1990, 2010, and Artstein 2002).

- (15) a. *Ite est leghende (Giuanne)?*
 what is reading Giuanne
 “What is John/he reading?”
 b. *Su giornale est leghende.*
 the newspaper is reading
 “He is reading the newspaper.” (Jones 2013: 80)

In his description of the relevant examples, Jones maintains the generalization that *wh*-phrases do not normally bear the NPA, but offers no specific details regarding the actual position of the NPA in *wh*-questions. Even when he does so by using bold to indicate the word that bears the NPA, as in (16), it remains unclear whether we are dealing with the V pattern – with the NPA on the lexical verb – or with the R pattern, with the NPA on the rightmost constituent. Given that the lexical verb is in the rightmost position of the sentence, it is virtually impossible to distinguish between the two patterns in this example, and the prosodic pattern thus remains ambiguous:

- (16) *A chie as telefonadu?*
 to whom have.2SG telephoned
 “Who did you telephone?” (Jones 2013: 88)

This issue is addressed in the thorough overview of Sardinian intonational phonology by Vanrell et al. (2015). Based on an analysis of their own collected data, Vanrell et al. (2015) conclude that Sardinian presents either the WH or the R pattern, depending on the length of the *wh*-question: “In Sardinian we find a situation very similar to the one described by Ladd (2008) for Romanian, in that the most prominent accent in *wh*-questions is on the *wh*-word when the sentence is short, but later if the sentence is longer... When the utterance is longer (when the *wh*-question contains more than two accentable syllables), the nuclear accent moves on the last prosodic word” (Vanrell et al. 2015: 337).⁶ Their example of a longer *wh*-question (17), however, runs into the same problems as Jones’s example in (16): the *wh*-question ends with a verb and it is therefore virtually impossible to distinguish between the V and the R pattern:

- (17) *Itte li potto comprare?*
 what him/her.DAT can.1SG buy
 “What can I buy for him/her”

In order to overcome these difficulties, I designed a pilot prosodic experiment in which the tested stimuli contained a postverbal constituent. This experiment and the relevant results are discussed in the next section.

6. Note that Ladd (2008) cited in the quote is simply the second edition of Ladd (1996).

4. Sardinian *wh*-questions: The prosodic experiment

To examine the placement of the NPA in Sardinian direct *wh*-questions, I carried out a pilot production experiment. Eight native speakers of Logudorese Sardinian took part in this experiment (6 men and 2 women, from 29 to 51 years old).⁷ The experiment consisted of a reading task and was specifically designed to investigate the distribution of the NPA in *wh*-questions with a bare *wh*-phrase. For this reason, all experimental stimuli included a postverbal constituent, that is, a direct or indirect object in the order Wh+V+XP. This allowed us to unambiguously distinguish between the three typological patterns identified in the previous section: (i) NPA on *wh*-element (WH), (ii) NPA on the verb (V), and (iii) NPA on the rightmost constituent (R).

4.1 The experimental material and procedure

The experimental material consisted of 10 stimuli and, more precisely, of direct *wh*-questions with a transitive or ditransitive verb. In the stimuli the *wh*-element corresponded either to the subject (*chie* “who”) or to the indirect object (*a chie* “to whom”): this was specifically designed so that in each stimulus the postverbal constituent was the direct or indirect object, that is, a syntactically integrated argument of the verb. Examples are given in (18):

- (18) a. *Chie at nadu chi as a cramare a Jubanne?*
 who has said that have.PRS.2SG to call to Jubanne
 “Who said you will call Jubanne?” (Pattada)
- b. *A chie as pedidu unu passàggiu?*
 to whom have.PRS.2SG asked a ride
 “Who did you ask for a ride?” (Pattada)

Participants were asked to produce each stimulus twice. Speakers did not receive any kind of feedback concerning the reading task. In some cases, however, they spontaneously asked to repeat a sentence in case of breaks or other kinds of disfluency. In these cases, they were allowed to repeat the sentence and the disfluent productions were discarded. A total of 160 target sentences were thus collected, segmented, and ToBI-transcribed: 8 speakers * 10 items * 2 (disfluency-free) repetitions.

In the transcriptions and annotations, the NPA was labelled as the rightmost PA after which the pitch contour is completely compressed and no fully-fledged PA is observable. To control for information-structure effects, the target sentences were

7. The exact origins of the speakers are indicated in Table 1 below.

introduced by a brief context eliciting a neutral question. In particular, a ‘given’ interpretation of the postverbal constituent was intentionally avoided to prevent this constituent from being construed (and hence produced) as dislocated to the right.

4.2 Results

The distribution of the NPA observed in our data is illustrated in Figure 1:

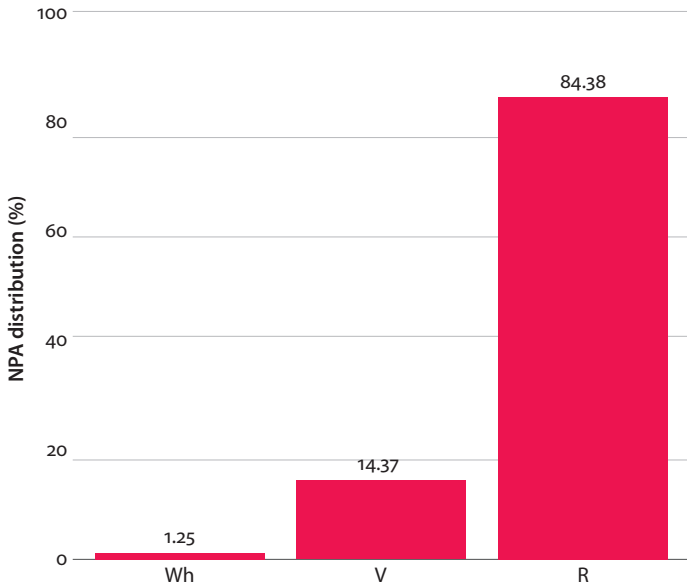


Figure 1. Prosodic experiment: NPA distribution in Sardinian *wh*-questions

As is evident from Figure 1, in the majority of the cases (84.38%), the NPA falls on the rightmost constituent of the sentence (i.e., the R pattern). This means that with respect to the prosody of (direct) *wh*-questions Sardinian patterns together with Western Romance languages such as Spanish, Catalan, and Portuguese. Figures 2 and 3 illustrate the pitch contour of a sentence produced with this pattern in two Sardinian varieties, namely the dialects of Pattada (19) and Fonni (20), respectively:

(19) *A chie as pedidu unu passàggiu?* (Pattada)
 to whom have.PRS.2SG asked a ride
 “Who did you ask for a ride?”

(20) *A chie as pediu unu passàggiu?* (Fonni)
 to whom have.PRS.2SG asked a ride
 “Who did you ask for a ride?”

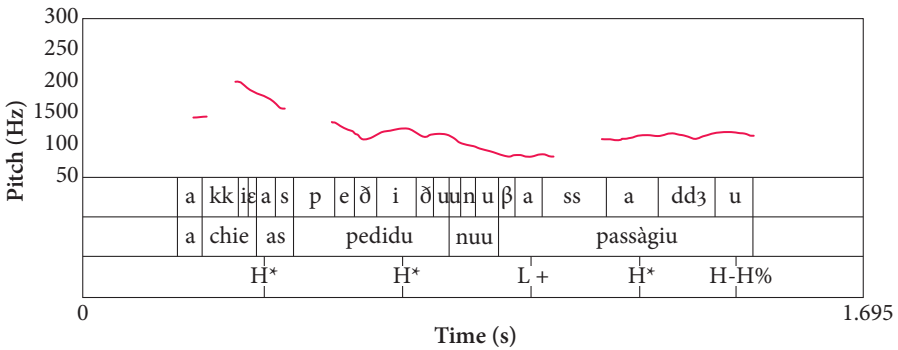


Figure 2. Prosodic experiment: Pitch contour of an utterance produced for (19) by speaker SB(M) from Pattada

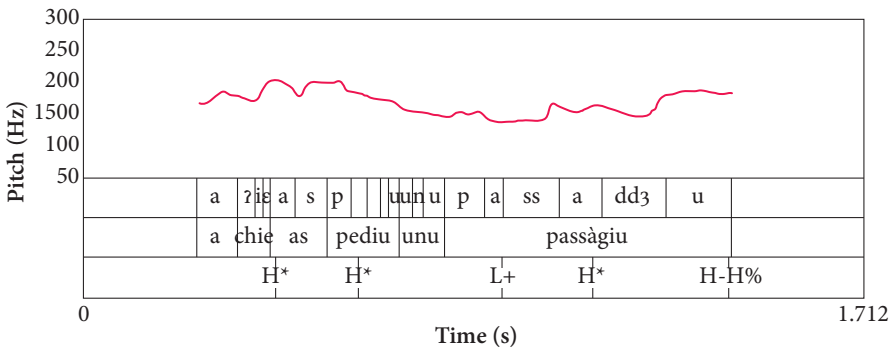


Figure 3. Prosodic experiment: Pitch contour of an utterance produced for (20) by speaker MM(F) from Fonni

Both in Figure 2 and Figure 3, an L+H* NPA aligns with the stressed syllable of the rightmost constituent of the *wh*-question (*passàgiu*), as expected in the R pattern.

Only in a very marginal number of cases (1.25%) does the NPA associate with the *wh*-element;⁸ the association between the NPA and the lexical verb, on the other hand, does not prove entirely negligible, in that it accounts for 14.37% of cases. In

8. Our data do not therefore support the existence of the WH pattern for short questions in Sardinian, as claimed in Vanrell et al. (2015) (cf. § 3.2), even though it should be noted that our stimuli do not qualify as short questions according to their definition. In order to be able to distinguish between the V and the R pattern, our experimental stimuli contained a postverbal constituent. Given that the experiment did not include short questions, we were thus not able to directly test Vanrell et al.'s (2015) claim that the prosodic pattern in Sardinian depends on the length of the *wh*-question. Some preliminary investigation with native speakers, however, appears to contradict this claim.

order to understand this last divergence with respect to what can be considered the general pattern of Sardinian *wh*-questions (i.e., the R pattern), we need to look at the results for individual speakers, given in Table 1:⁹

Table 1. The prosodic experiment: Results for individual speaker

Speakers			NPA distribution		
ID	ORIGIN	AGE	Wh	V	R
SB(M)	Orani	30		20 (100%)	
SF(M)	Pattada	33			20 (100%)
GC(M)	Berchidda	33			20 (100%)
SFa(M)	Ittiri	29		3 (15%)	17 (85%)
RF(M)	Oliena	29			20 (100%)
FP(M)	Oliena	51	1 (5%)		19 (95%)
AG(F)	Oliena	50	1 (5%)		19 (95%)
MM(F)	Fonni	40			20 (100%)
			2 (1.250%)	23 (14.375%)	135 (84.375%)

This table clearly shows that only one speaker is responsible for the divergent pattern, that is, the V pattern: SB(M) from Orani consistently produced the V pattern in 100% of the stimuli. The difference is clear if we compare the pitch contour in Figure 4, produced by SB(M) for the sentence in (20), with the pitch contours in Figures 2 and 3:

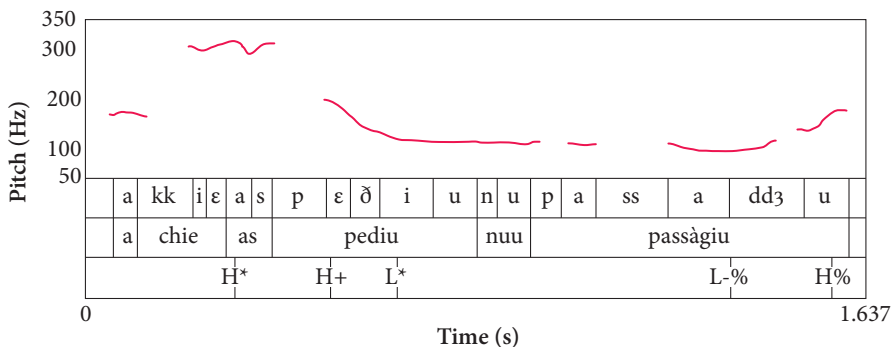


Figure 4. Prosodic experiment: Pitch contour of an utterance produced for (20) by speaker SB(M) from Orani

9. The ID of each speaker consists of their initials followed by an indication of their gender in brackets.

In Figure 4, we see that the H+L* NPA does not align with the stressed syllable of the rightmost constituent, as expected in the R pattern, but falls instead on the lexical verb (*pediu*). After this NPA, the pitch contour is compressed and no fully-fledged PAs are observable. This contour corresponds to the V pattern typical of Italian. The identity of the patterns becomes evident if we compare Figure 4 with the pitch contour of an Italian *wh*-question such as the one illustrated in Figure 5 and produced for the sentence in (21) by a Tuscan speaker (from Bocci et al. 2021):

- (21) *A chi hanno chiesto un autografo dopo lo spettacolo?* (Italian)
 to whom have.3PL asked an autograph after the show
 “Who did they ask for an autograph after the show?”

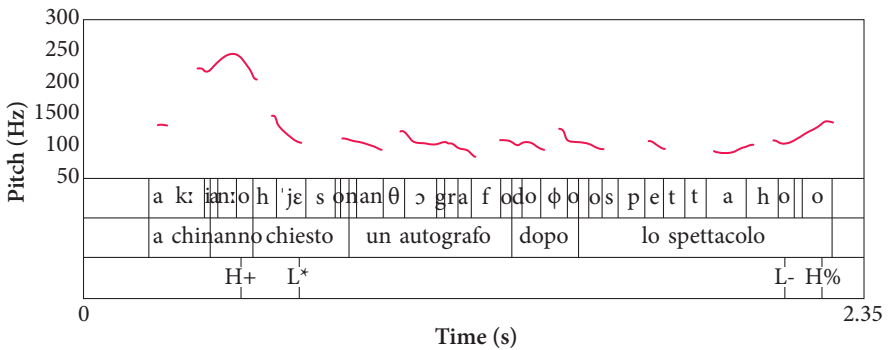


Figure 5. Prosodic experiment: Pitch contour of an utterance produced for (21) by a Tuscan speaker from Siena

Although these findings require further examination and testing, possibly with the additional aid of an accurate analysis of sociolinguistic variables (a task that I leave to future research), we can hypothesize that the V pattern of the speaker from Orani reflects a case of prosody in a language-contact situation, where two language systems (i.e., the Sardinian R pattern and the Italian V pattern) co-exist. Like most native Sardinian speakers, SB(M) is indeed a bilingual speaker with both Sardinian and Italian as L1. This kind of transfer or influence at the prosodic level in the context of language contact is not in fact an isolated case: similar effects of phonological and prosodic influence have been discussed with reference to other bilinguals, both with respect to production and perception (see, e.g., Sichel-Bazin, Buthke & Meisenburg 2012; Vanrell et al. 2015; Vanrell & Fernández-Soriano 2013, 2018; Muntendam & Torreira 2016).

5. The asymmetry between direct and indirect *wh*-questions

In § 2.1, we saw that FF and *wh*-movement are mutually exclusive in direct *wh*-questions, in that they cannot occur simultaneously within the same sentence, regardless of the respective ordering of the *wh*-phrase and the fronted focus. I repeat the relevant Sardinian examples here in (22):

- (22) a. **A kie su jocátulu as datu?*
 to whom the toy have.PRS.2SG given
 b. **Su jocátulu a kie as datu?*
 the toy to whom have.PRS.2sg given
 “To whom did you give the toy?”

Interestingly, the same incompatibility does not hold in indirect *wh*-questions, as shown in (23), a sentence with an embedded, indirect *wh*-question that could be uttered in a marked context, for example to correct a previous statement such as *I am sure you are wondering what they said to Paulu*:

- (23) *Mi so dimandande a Juanne itte appan natu (,no a Paulu).*
 Me be.1SG ask.GER to Juanne what had.3PL said not to Paulu
 “I’m wondering what they said to Juanne (, not to Paulu).”

The asymmetry between direct and indirect *wh*-questions with respect to the compatibility between *wh*-phrases and foci was first noted by Rizzi (2001), who shows that the co-occurrence of focus and *wh*-phrase in the same clause, which is excluded in main questions (24), becomes (at least marginally) possible in indirect embedded questions (25) (see also Bocci et al. 2018 for experimental evidence):¹⁰

- (24) a. **A chi questo hanno detto (non qualcos’ altro)?*
 to whom this have.3PL said not something else
 b. **Questo a chi hanno detto (non qualcos’ altro)?*
 this to whom have.3PL said not something else
 c. **A Gianni che cosa hanno detto (non a Piero)?*
 to Gianni what have.3PL said not to Piero
 d. **Che cosa a Gianni hanno detto (non a Piero)?*
 what to Gianni have.3PL said not to Piero

(Rizzi 2001: 290–291)

- (25) *Mi domando a Gianni che cosa abbiamo detto (non a Piero).*
 me ask.1SG to Gianni what have.SBJV.3PL said not to Piero
 “I wonder what they said to Gianni (not to Piero).” (Rizzi 2001: 291)

10. The *wh*-element *perché* “why” is also exceptional with respect to this property (see Rizzi 2001; Bianchi et al. 2017). See footnote 3.

How can we account for this asymmetry? Two hypotheses present themselves. First, we could argue that *wh*-words are focal in direct questions, but not in indirect questions. Alternatively, as a second hypothesis, we could claim that *wh*-words do not have focal status in any syntactic context, i.e., they are not inherently focal either in direct or in indirect *wh*-questions. The first hypothesis seems to be rather strong, inasmuch as it presupposes that the compositional semantics of direct *wh*-questions is radically different from that of indirect questions. The second hypothesis, by contrast, would more directly explain the prosodic facts, namely the lack of alignment between the NPA and the *wh*-phrase in several languages including Italian and Sardinian, but it requires an additional explanation for the root-embedded asymmetry. It should also be noted that in indirect questions *wh*-words are typically unstressed, but unlike in direct questions, they need not be adjacent to the verb and can be followed by a preverbal subject, as shown in (26) from Sardinian (see Bocci & Pozzan 2014; Bocci & Cruschina 2018; and Bocci et al. 2018 on Italian):

- (26) *No isco in ube Juanne at istikkitu su dinari.*
 not know.1SG where Juanne has hidden the money
 “I do not know where John has hidden the money.” (Jones 1993: 336)

If we adopt the second hypothesis, such that the main semantic and prosodic properties of *wh*-questions are considered to be the same in both direct and indirect questions, we then need to account for the difference in their syntactic distribution. Drawing on Bianchi et al. (2017), I follow this hypothesis and propose that *wh*-phrases are not inherently focal, but that while indirect questions enjoy a free focal structure, the focal structure of direct questions is independently more constrained.

6. Analysis: Interrogative *wh*-words are not inherently focal

Direct *wh*-questions require a congruent answer. Standard analyses of focus define congruence in terms of identity between the question meaning and the focal alternatives of the answer, so that felicitous and ‘congruent’ answers must contain a narrow focus on the constituent that corresponds to the *wh*-phrase in the question (see, a.o., Rooth 1992; Roberts 1996, 2012; Krifka 2001, 2011; Reich 2002):

- (27) a. [*What*] *did Ed drink?*
 b. *Ed drank [coffee]_F.*
- (28) a. [*Who*] *drank coffee?*
 b. [*Ed*]_F *drank coffee.*

Sardinian and other languages that display the R pattern: unlike the [focus]-marked phrase that moves in the case of FF, the *wh*-phrase of a *wh*-question does not qualify for the assignment of the NPA by the phonological component. Consequently, ‘default’ assignment – the same as in declaratives – take place, yielding the association between the NPA and the stressed syllable of the rightmost constituent (on the V pattern in Italian, see Bocci et al. 2021). In other words, at the prosodic level, no element within a *wh*-question is interpreted as narrow focus and the sentence is treated as an instance of broad focus lacking a special focal marking, along the lines of unmarked declarative clauses.

7. Conclusions

On the basis of evidence from prosody and the different syntactic behaviour of indirect questions, particularly in Sardinian, I have argued in this paper that *wh*-words are not inherently focal. The syntactic similarities between FF and *wh*-movement and the fact that they are mutually exclusive can be explained by the interpretive requirement that no focus should be allowed in direct *wh*-questions in order to yield a congruent answer. This constraint can be implemented with the long-standing idea that FF and *wh*-movement target the same projection, and with the additional assumption that the features [wh] and [focus] are bundled together in the same projection. In indirect questions, the features [wh] and [focus] are not necessarily bundled together and can be checked independently in different projections by distinct phrases: hence, FF and *wh*-movement can co-occur. The lack of a [focus]-marked element in *wh*-questions explains why the *wh*-phrase is not interpreted as focal at the phonological level and therefore, in several languages, including Sardinian, is not associated with the NPA of the sentence: prosodically, focus needs to be maximally prominent; since *wh*-phrases are not maximally prominent, they are not inherently focal.

These results raise additional questions and suggest fruitful directions for further research. First, we need to examine the extent of the prosodic influence of the Italian V pattern on the Sardinian R pattern, which is limited to one speaker in our data. A more extensive investigation of this potential effect of language contact could provide crucial insights into language change – presumably, a change in progress – at the level of intonational phonology and in the domain of bilingualism. Secondly, while the R pattern supports the proposal that *wh*-phrases are not inherently focal, a more careful comparison with the other patterns is required in order to arrive at a theoretical explanation of the parametric variation at issue. Indeed, it remains possible that in some languages *wh*-phrases are prosodically treated as focal, as should be the case in languages that exhibit the WH pattern with respect

to the assignment of the NPA in *wh*-questions (Romanian, Hungarian, and Greek; cf. § 3.1).¹² Finally, the relation between the two features [focus] and [wh] should also be investigated with respect to their role in and contribution to the overall syntax and semantics of questions, both in direct and indirect clauses (see, e.g., Aboh & Pfau 2010).

In conclusion, we need to investigate how the empirical richness of the prosodic variation in *wh*-questions can be captured in a coherent and principled theoretical interface model. With this paper, I hope to have contributed a small step in this direction.

Acknowledgements

I would like to thank Valentina Bianchi for helpful comments on some theoretical aspects of this paper, and Giuliano Bocci for help with the prosodic analysis. I am also grateful to Stefano Fogarizzu and Simone Pisano for their assistance with the data collection and the recordings.

References

- Aboh, Enoch O. & Roland Pfau. 2010. What's a *wh*-word got to do with it? In Paola Benincà & Nicola Munaro (eds.), *Mapping the left periphery. The cartography of syntactic structures*, Volume 5, 91–124. Oxford: Oxford University Press.
- Alexopoulou, Theodora & Mary Baltazani. 2012. Focus in Greek *wh*-questions. In Ivona Kučerová & Ad Neeleman (eds.), *Information structure: Contrasts and positions*, 206–246. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511740084.011>
- Artstein, Ron. 2002. Parts of words: Compositional semantics for prosodic constituents. Ph.D. Dissertation: Rutgers University.
- Beck, Sigrid. 2006. Intervention effects follow from focus interpretation. *Natural Language Semantics* 14. 1–56. <https://doi.org/10.1007/s11050-005-4532-y>
- Beckman, Mary E., Julia Hirschberg & Stefanie Shattuck-Hufnagel. 2005. The original ToBI system and the evolution of the ToBI framework. In Sun- Ah Jun (ed.), *Prosodic typology: The phonology of intonation and phrasing*, 9–54. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199249633.003.0002>
- Bianchi, Valentina, Giuliano Bocci & Silvio Cruschina. 2017. Two types of subject inversion in Italian *wh*-questions. *Revue Roumaine de Linguistique* 62(3). 233–252.

12. It must be noted that even in these languages – at least in some of them – the root-embedded asymmetry still holds not only with respect to the compatibility between *wh*-phrases and fronted foci (cf. § 5), but also with respect to the prosodic pattern of indirect *wh*-questions. See Alexopoulou & Baltazani (2012) for Greek.

- Bocci, Giuliano & Silvio Cruschina. 2018. Postverbal subjects and nuclear pitch accent in Italian *wh*-questions. In Roberto Petrosino, Pietro Cerrone & Harry van der Hulst (eds.), *From sounds to structures. Beyond the veil of Maya*, 467–494. Berlin: De Gruyter.
<https://doi.org/10.1515/9781501506734-017>
- Bocci, Giuliano & Lucia Pozzan. 2014. Questions (and experimental answers) about Italian subjects. Subject positions in main and indirect question in L1 and attrition. In Carla Contemori & Lena Dal Pozzo (eds.), *Inquiries into linguistic theory and language acquisition. Papers offered to Adriana Belletti*, 28–44. Siena: CISCL Press.
- Bocci, Giuliano, Luigi Rizzi & Mamoru Saito. 2018. On the incompatibility of *wh* and focus. *Gengo Kenkyui – Journal of the Linguistic Society of Japan* 154. 29–51.
- Bocci, Giuliano, Valentina Bianchi & Silvio Cruschina. 2021. Focus in *wh*-questions: Evidence from Italian. *Natural Language & Linguistic Theory* 39. 405–455.
<https://doi.org/10.1007/s11049-020-09483-x>
- Bocci, Giuliano, Silvio Cruschina & Luigi Rizzi. 2021. On some special properties of *why* in syntax and prosody. In Gabriela Soare (ed.), *Why is 'why' unique? Its syntactic and semantic properties*, 293–316. Berlin: De Gruyter.
- Brody, Michael. 1995. Focus and checking theory. In Istvan Kenesei (ed.), *Approaches to Hungarian 5: Levels and structures*, 29–43. Szeged: JATEPress.
- Büring, Daniel. 2003. On D-trees, beans, and B-accent. *Linguistics and Philosophy* 26. 511–545.
<https://doi.org/10.1023/A:1025887707652>
- Cable, Seth. 2010. *The grammar of Q: Q-particles, wh-movement, and pied-piping*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195392265.001.0001>
- Chomsky, Noam. 1976. Conditions on rules of grammar. *Linguistic Analysis* 2. 303–352.
- Eckardt, Regine. 2007. Inherent focus on *wh*-phrases. *Proceedings of Sinn und Bedeutung* 11. 209–228.
- Frota, Sónia, Marisa Cruz, Flaviane Svartman, Gisela Collischonn, Aline Fonseca, Carolina Serra, Pedro Oliveira & Marina Vigário. 2015. Intonational variation in Portuguese: European and Brazilian varieties. In Sónia Frota & Pilar Prieto (eds.), *Intonational variation in Romance*, 235–283. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199685332.003.0007>
- Haegeman, Liliane & Jacqueline Guéron. 1999. *English grammar: A generative perspective*. Malden: Blackwell.
- Hualde, José Ignacio & Pilar Prieto. 2015. Intonational variation in Spanish: European and American varieties. In Sónia Frota & Pilar Prieto (eds.), *Intonational variation in Romance*, 350–391. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199685332.003.0010>
- Jones, Michael A. 1993. *Sardinian syntax*. London: Routledge.
- Jones, Michael A. 2013. Fronting, focus and illocutionary force in Sardinian. *Lingua* 134. 75–101.
<https://doi.org/10.1016/j.lingua.2013.07.014>
- Krifka, Manfred. 2001. For a structured account of questions and answers. In Caroline Féry & Wolfgang Sternefeld (eds.), *Audiatur vox sapientiae. A festschrift for Arnim von Stechow*, 287–319. Berlin: Akademie-Verlag. <https://doi.org/10.1515/9783050080116.287>
- Krifka, Manfred. 2008. Basic notions of information structure. *Acta Linguistica Hungarica* 55. 243–276. <https://doi.org/10.1556/ALing.55.2008.3-4.2>
- Krifka, Manfred. 2011. Questions. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics. An international handbook of natural language meaning*, Volume 2, 1742–1758. Berlin: Mouton de Gruyter.

- Krifka, Manfred. Manuscript. *Focus and contrastive topics in question and answer acts*. Humboldt-Universität Berlin & Zentrum für Allgemeine Sprachwissenschaft (ZAS) Berlin. 2017.
- Ladd, D. Robert. 1996. *Intonational phonology*. Cambridge: Cambridge University Press (Second edn. in 2008).
- Marotta, Giovanna. 2001. I toni accentuali nelle interrogative aperte (wh-) dell'italiano di Lucca. In Camilla Bettoni, Antonio Zampolli & Daniela Zorzi (eds.), *Atti del II congresso di studi dell'Associazione Italiana di linguistica applicata*, 175–194. Perugia: Guerra Edizioni.
- Muntendam, Antje & Francisco Torreira. 2016. Focus and prosody in Spanish and Quechua: Insights from an interactive task. In Meghan E. Armstrong, Nicholas Hendriksen & Maria del Mar Vanrell (eds.), *Intonational grammar in Ibero-Romance: Approaches across linguistic subfields*, 69–90. Amsterdam: John Benjamins. <https://doi.org/10.1075/ihill.6.04mun>
- Prieto, Pilar, Joan Borràs-Comes, Teresa Cabré, Verònica Crespo-Sendra, Ignasi Mascaró, Paolo Roseano, Rafèu Sichel-Bazin & Maria del Mar Vanrell. 2015. Intonational phonology of Catalan and its dialectal varieties. In Sónia Frota & Pilar Prieto (eds.), *Intonational variation in Romance*, 9–62. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199685332.003.0002>
- Reich, Ingo. 2002. Question/answer congruence and the semantics of wh-phrases. *Theoretical Linguistics* 28. 73–94. <https://doi.org/10.1515/thli.2002.28.1.73>
- Reis, Marga. 1991. Echo-w-Sätze und Echo-w-Fragen. In Marga Reis & Inger Rosengren (eds.), *Fragesätze und Fragen. Referate Anlässlich der 12. Jahrestagung der Deutschen Gesellschaft für Sprachwissenschaft, Saarbrücken 1990*, 49–76. Tübingen: Niemeyer.
<https://doi.org/10.1515/978311356525.49>
- Reis, Marga. 2012. On the analysis of echo questions. *Tampa Papers in Linguistics* 3. 1–24.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar: Handbook of generative Syntax*, 281–337. Dordrecht: Kluwer.
https://doi.org/10.1007/978-94-011-5420-8_7
- Rizzi, Luigi. 2001. On the position “Int(errogative)” in the left periphery of the clause. In Guglielmo Cinque & Giampaolo Salvi (eds.), *Current studies in Italian syntax. Essays offered to Lorenzo Renzi*, 287–296. Amsterdam: Elsevier.
- Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. In Jae-Hak Yoon & Andreas Kathol (eds.), *Ohio State University Working Papers in Linguistics 49. Papers in Semantics*, 91–136. Columbus: The Ohio State University.
- Roberts, Craige. 2012. Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics* 5. 1–69. <https://doi.org/10.3765/sp.5.6>
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1. 75–116.
<https://doi.org/10.1007/BF02342617>
- Sichel-Bazin, Rafèu, Carolin Buthke & Trudel Meisenburg. 2012. The prosody of Occitan-French bilinguals. In Kurt Braunmüller & Christoph Gabriel (eds.), *Multilingual individuals and multilingual societies*, 349–364. Amsterdam: John Benjamins.
<https://doi.org/10.1075/hsm.13.23baz>
- Sobin, Nicholas. 1990. On the syntax of English echo questions. *Lingua* 81. 141–167.
[https://doi.org/10.1016/0024-3841\(90\)90010-l](https://doi.org/10.1016/0024-3841(90)90010-l)
- Sobin, Nicholas. 2010. Echo questions and the Minimalist Program. *Linguistic Inquiry* 41. 131–148.
<https://doi.org/10.1162/ling.2010.41.1.131>
- Truckenbrodt, Hubert. 1995. Phonological phrases: Their relation to syntax, focus, and Prominence. Ph.D. Dissertation: Massachusetts Institute of Technology.

- Vanrell, Maria del Mar & Olga Fernández-Soriano. 2013. Variation at the interfaces in Ibero-Romance. Catalan and Spanish prosody and word order. *Catalan Journal of Linguistics* 12. 253–282. <https://doi.org/10.5565/rev/catjl.63>
- Vanrell, Maria del Mar & Olga Fernández-Soriano. 2018. Language variation at the prosody-syntax interface. Focus in European Spanish. In Marco García García & Melanie Uth (eds.), *Focus realization in Romance and beyond*, 33–70. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.201.02van>
- Vanrell, Maria del Mar, Francesc Ballone, Carlo Schirru & Pilar Prieto. 2015. Sardinian intonational phonology: Logudorese and Campidanese varieties. In Sónia Frota & Pilar Prieto (eds.), *Intonational variation in Romance*, 317–349. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199685332.003.0009>
- Zubizarreta, María-Luisa. 1998. *Prosody, focus, and word order*. Cambridge, MA: MIT Press.
- Zubizarreta, María-Luisa. 2016. Nuclear stress and information structure. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 165–184. Oxford: Oxford University Press.
- Zubizarreta, María-Luisa & Jean-Roger Vergnaud. 2005. Phrasal stress, focus, and syntax. In Martin Everaert & Henk van Riemsdijk (eds.), *The syntax companion*, 522–556. Oxford: Blackwell.

The varieties of temporal anaphora and temporal coincidence

Hamida Demirdache

Nantes University / CNRS

This paper explores the temporal construals of perfective vs. imperfective aspect in Sequence of Tense contexts in Spanish and French, in particular, under ellipsis. The distribution of past-shifted vs. simultaneous, as well as sloppy vs. strict, temporal construals is taken to support extending to viewpoint aspect a referential approach to tense, as Demirdache & Uribe-Etxebarria (2014) contend. I derive the distribution of simultaneous vs. past-shifted readings by extending their analysis of imperfective vs. perfective to embedded contexts. The intricate distribution of strict vs. sloppy (simultaneous, as well as past-shifted) readings is explained by extending the LF-parallelism constraint on ellipsis (Fox 2000) – specifically, the assumption that structural parallelism yields sloppy readings, while referential parallelism yields strict readings – to temporal anaphora under ellipsis.

Keywords: perfective, imperfective, sequence of tense, simultaneous past, past-shifting, temporal ellipsis, temporal binding, strict past, sloppy past, temporal parallelism

1. Introduction

This paper adduces novel evidence from Sequence of Tense (SoT) contexts for Demirdache & Uribe-Etxebarria's (2014, 2007, and references therein; henceforth D&UE) contention that the parallels established in the literature between tenses and pronouns (Partee 1973; Kratzer 1998 a.o.) should be extended to the domain of viewpoint aspect. Assuming a close parallel between anaphora in the nominal and temporal domains, together with a model of temporal interpretation where aspect, just like tense, serves to relate two times (Klein 1994), D&UE seek to extend the referential approach developed for tense to the domain of viewpoint aspect. The starting tenet is that anaphora between individual-denoting DPs can be resolved as either (variable) binding or coreference, distinguishable in the appropriate context (Reinhart 1997; Heim 1998). The null assumption on a referential approach to both

tense and aspect is that anaphora between the time-denoting DPs/Zeit-Ps projected in the syntax as arguments of T° , Asp° , V° , can also be resolved as either binding or coreference, likewise distinguishable in the appropriate context. The specific proposal defended by D&UE (2014) is that coreference yields a perfective viewpoint, while (variable) binding yields an imperfective viewpoint. This hypothesis is presented in § 2, in the context of the general framework for temporal relations developed in D&UE.

I extend here D&UE's (2014) analysis for (im)perfective past to SoT contexts. § 3 first shows how this hypothesis accounts for one of the most well-known differences between imperfective and perfective in subordinate contexts: while an imperfective past (the *imparfait* in French or the *imperfecto* in Spanish) can be used to convey either a simultaneous or a past-shifted reading relative to a matrix past, a perfective past only allows the past-shifted reading. In line with recent accounts of SoT phenomena (Altshuler & Schwarzschild 2012; Altshuler 2016; Kauf & Zeijlstra 2018), the proposal developed here does not posit tense deletion/zero-tense to account for the simultaneous reading. Rather, on both readings, the embedded past is semantically interpreted, but while an imperfective viewpoint – established via temporal binding – yields either a simultaneous or a past shifted construal (§ 3.1), a perfective viewpoint – established via temporal coreference – only yields the past shifted reading (§ 3.2).

A standard diagnostic in the literature on anaphora for distinguishing between coreference and Bound Variable Anaphora (BVA) is VP ellipsis, on the classic assumption that coreference yields strict identity readings under ellipsis, while BVA yields sloppy identity readings (Sag 1976). Building on Stowell's (2014) commentary on D&UE (2014), § 4 probes the interpretations of imperfective vs. perfective aspect in SoT contexts under VP/TP ellipsis, identifying an intricate pattern of distribution of sloppy vs. strict readings. I derive this pattern by adapting the LF-Parallelism constraint on nominal ellipsis (Fox 2000) to temporal ellipsis, and in so doing extending the proposal that strict readings involve referential parallelism, while sloppy readings involve structural parallelism to temporal anaphora under ellipsis. § 5 concludes.

2. Temporal anaphora and viewpoint aspect

Let's start by succinctly presenting the relevant aspects of D&UE's framework. In line with Stowell (2007), the approach developed in D&UE incorporates insights from both 'referential' theories of tense (Partee 1973; Enç 1987; Kratzer 1998, a.o.) and 'predicative' theories of tense (Comrie 1976). It shares with predicative theories the assumption that tenses are predicates of temporal ordering, while extending this assumption to aspect, and with referential approaches, the assumption that the

semantics of tense involve pronominal like expressions referring to times, projected in the syntax as time-denoting DPs (Zeit-Ps, for Stowell (2007)), again extending this approach to aspect.

2.1 Temporal syntax of tense & aspect

D&UE uniformly analyze tenses, aspects and time adverbs as predicates of spatio-temporal ordering taking time-denoting arguments, as the typology below illustrates.

Table 1. Spatiotemporal predicates

	AFTER Subsequence	WITHIN Inclusion	BEFORE Precedence
Tense	Past	Present	Future
Aspect	Perfect	Progressive	Prospective
Locating adverbs	<i>after</i> DP/CP	<i>at, in, during</i> DP, <i>when</i> CP	<i>before</i> DP/CP
Durational adverbs	<i>since</i> DP/CP	<i>during</i> DP, <i>while</i> CP	<i>until</i> DP/CP

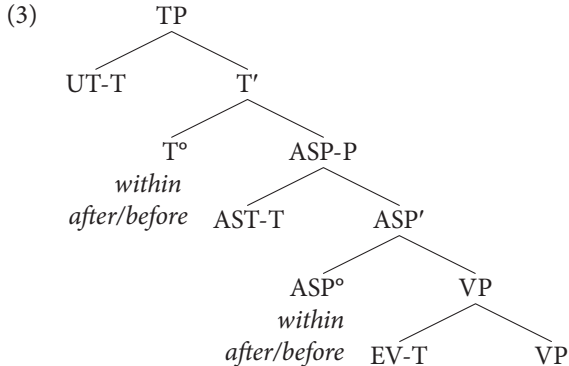
Viewpoint aspect focuses an interval in the temporal contour of the described event. This interval is the Assertion-Time: “the time for which an assertion is made ... the time to which the assertion made by the utterance is confined” (Klein 2000: 366). Aspect focuses an interval within the time span defining the described event by ordering the time of the assertion (AST-T) relative to the event-time (EV-T). If this ordering relation is subsequence, then the resulting viewpoint corresponds to retrospective aspect (that is, perfect aspect, cf. the discussion of (4c)). If this relation is inclusion, then the viewpoint corresponds to progressive aspect, and if it is precedence, to prospective aspect, as illustrated with the schemas in (1).

- (1) a. Retrospective
 AST-T AFTER EV-T
 EV-T AST-T
 —[—]—[—]—>
- b. Progressive
 AST-T WITHIN EV-T
 AST-T
 —[—[—]]—>
 EV-T
- c. Prospective
 AST-T BEFORE EV-T
 AST-T EV-T
 —[—]—[—]—>

Tense, just like aspect, is a spatiotemporal predicate relating two times: the AST-T relative to another REF-T – the UT-T in matrix clauses. Again, the ordering relation can be one of subsequence, inclusion, or precedence:

- (2) a. Past
 UT-T AFTER AST -T
 AST -T UT-T
 —[——]—[——]—>
- b. Present
 UT-T WITHIN AST -T
 UT-T
 —[——[——]—]—>
 AST-T
- c. Future
 UT-T BEFORE AST-T
 UT-T AST-T
 —[——]—[——]—>

Time arguments are mapped onto a universal hierarchy of functional projections (TP < ASP-P < VP), as shown in (3).



2.2 More viewpoints: Imperfective and perfective

With the proposal outlined above, D&UE seek to constrain the set of logically possible tenses and aspects, while at the same time predicting the striking diversity of predicates that surface across languages to express temporal relations. Importantly, this typology is not random but systematic and, as such, predictable: pre/postpositions such as *in*, *on*, *at* as well as verbs of location, stance or posture (e.g., *stay*,

sit, lie, stand, live) or even verbs of non-directed motion (e.g., *walk*) are used to express inclusion (i.e., Progressive/Durative), while predicates of motion towards the ground (prepositions such as *to(wards)* or verbs such as *go*) are used to express precedence (i.e., Prospective/Future), and predicates of motion from the ground (prepositions such as *after* or verbs such as *leave, come from* or *throw away*) are used to express subsequence (i.e., Perfect/Past).

As D&UE (2014) argue, however, the proposal as summarized in § 2.1 is too restrictive, in that it fails to derive two other viewpoints: perfective and imperfective. To understand why, consider the denotations given by Kratzer (1998: 17) for these viewpoints.

- (4) a. Imperfective: reference-time included in event-time
 $\lambda P_{\langle \langle s, t \rangle \rangle} . \lambda t_i . \lambda w_s . \exists e_1 (t \subseteq \text{time}(e) \ \& \ P(e)(w) = 1)$
 b. Perfective: event-time included in reference-time
 $\lambda P_{\langle \langle s, t \rangle \rangle} . \lambda t_i . \lambda w_s . \exists e_1 (\text{time}(e) \subseteq t \ \& \ P(e)(w) = 1)$
 c. Perfect: event-time over by reference-time
 $\lambda P_{\langle \langle s, t \rangle \rangle} . \lambda t_i . \lambda w_s . \exists e_1 (\text{time}(e) < t \ \& \ P(e)(w) = 1)$

Note that temporal subsequence in D&UE's framework yields what corresponds to perfect aspect for Kratzer in (4c) since, when ASP^o orders the AST-T after the EV-T, then the event-time is understood as over by a reference-time (the AST-T), as illustrated with the schema in (1a). The question is then how to integrate so-called perfective ('aoristic') aspect, the viewpoint under which the described event is portrayed in its entirety, from its initial to its final bound (cf. Smith 1991: 66). Perfective was derived in (4b) by requiring the event-time to fall within the reference-time. Crucially, however, under the restrictive assumption that the time arguments of tense and aspect are mapped/projected onto the universal hierarchy of functional projections TP < ASP-P < VP in (3), then inclusion cannot yield (be read off the phrase structure in (3) as) perfective (EV-T ⊂ REF-T). That is, ASP orders the Zeit-P in its specifier relative to the Zeit-P it immediately c-commands, ASP_{WITHIN} thus automatically yields the ordering AST-T/REF-T ⊂ EV-T, corresponding to Progressive aspect (1a) – and not EV-T ⊂ AST-T/REF-T, corresponding to Perfective (4b).

Notice, further, that inclusion of the REF-T within the EV-T has been taken to yield progressive aspect, and not imperfective aspect, contrary to Kratzer in (4a). Evidence for this move comes from critical distinctions between progressive and imperfective viewpoints in Romance.

The paradigms in (5)–(6) illustrate how the French and Spanish imperfective pasts, the so-called *Imparfait* and *Imperfecto* (henceforth *Imp*), interact with punctual temporal adverbs. (5) shows the expected pattern of temporal modification with the imperfective: the described activity is understood as ongoing at the REF-T

specified by the time adverbial. However, in contexts involving iteration of events (e.g., habitual as in (6)), modification of an activity/accomplishment in the *Imp* by a punctual adverb also yields the ‘sequential’ construal ((6ii)) characteristic of perfective viewpoint ((7)).

(5) *Imp* (ongoing) viewpoint

a. *Ayer, a las seis, Nora sonreía.*

b. *Hier, à six heures, Nora riait.*

Yesterday at six hours Nora laugh_{imp}

“Yesterday, at six, Nora was (in the process of) laughing.”

√Ongoing

(6) *Imp* (neutral) viewpoint

a. *Amina siempre escribía un poema a las seis*

Amina always write_{imp} a poem at the six

b. *Amina écrivait toujours un poème à six heures*

Amina write_{imp} always a poem at six hours

i. “Amina was always (in the process of) writing a poem at six.”

√Ongoing

ii. “Amina always wrote (started writing) a poem at six.”

√Sequential

(7) *Perfective* viewpoint

a. *Amina escribió un poema a las seis*

b. *Amina écrivit un poème à six heures*

Amina write_{perf} a poem at six hours

“Amina wrote (started writing) a poem at six.”

√Sequential

The availability of sequential construals is taken to show that the aspectual viewpoint of the *Imp* is not imperfective, but rather ‘neutral’, and as such compatible with both imperfective and perfective construals (Smith 1991). (8) from Kamp & Rohrer (1983: 259) illustrates another notorious perfective use of the *Imp*: it can advance narrative time forward, unlike an English past progressive. (9) shows that habitual construals of the *Imp* allow modification by completive (‘in *x*-time’) adverbials, a hallmark test for temporally bounded interpretations.

(8) *Jean tourna l'interrupteur. La lumière éclatante l'éblouissait*

Jean turn_{preterit} the light. The light bright him.blind_{imp}

“Jean switched on the light. The bright light blinded his eyes.”

(9) a. *Amina écrivait un poème en une demi-heure*

Amina write_{IMP} a poem in a half-hour

“Amina wrote (/used to write) a poem in half hour.”

b. *Amina escribía un poema en media hora*

Amina write_{IMP} a poem in half hour

“Amina wrote (/used to write) a poem in half hour.”

Importantly, the French (periphrastic) progressive *en train de*, the Spanish progressive (*estar + gerundio*), just like the English progressive, do not yield temporally bounded construals under these tests, as the unavailability of a sequential reading in (10) illustrates.

(10) *Progressive viewpoint*

- a. *Amina était toujours en train d'écrire un poème à six heures*
Amina be_{imp} always in process of write_{inf} a poem at six hours
- b. *Amina estaba siempre escribiendo un poema a las seis*
Amina be_{imp} always writing_{gerund} a poem at the six
“Amina was always writing a poem at six.” √ Ongoing, *Sequential

D&UE concur with Smith (1991), Laca (2005), or Schaden (2009) that the critical difference between the *Imp* and the progressive (at least as far as viewpoint is concerned) is that the progressive, unlike the *Imp*, excludes the initial and final bounds of the event from the viewpoint. The described event is thus viewed from ‘inside’, as an ongoing process. This is expected on the assumption that progressive orders the AST-T within the EV-T ($AST-T \subset EV-T$ (1b)). Since the bounds of the event are excluded from the AST-T, modification by the punctual adverb in (10) is correctly predicted to only yield an ongoing construal.

Summarising, two issues remain to be explained: how to integrate perfective vs. imperfective aspect, moreover distinguishing the imperfective from the progressive.

2.3 Temporal anaphora

Since aspect serves to provide a temporal viewpoint on the described event by ordering two times, the EV-T relative to a REF-T (our AST-T), D&UE (2014) incorporate (im)perfective aspect within their framework on the assumption that the ordering relation holding between these two times can be established via anaphora. Anaphora between individual-denoting DPs can be established via (variable) binding or coreference, which can be disambiguated in contexts such as focus ((11)), or ellipsis. On Reinhart’s (1997) proposal, when a variable is bound by an antecedent, there is predication over the antecedent with the variable bound by the predicate abstractor.¹ Binding is thus just “the procedure of closing a property” (Reinhart

1. Technically, the binder in (11b–d) is the λ -operator (predicate abstractor). Reinhart (1997: 6) captures the syntactic notion of binding with (i), which defines the DP *Amina* as the A-binder of the variable *x* in (11) (see also (15)).

(i) A-Binding (logical-syntax based definition):

a A-binds b iff a is the sister of a λ -predicate whose operator binds b.

1997: 6) (forming sets). Binding of the free variable y in (11b) yields a predicate denoting the set of female individuals who love their own father ((11c)). Under binding, the sentence thus asserts that only Amina belongs to this set – that is, other women do not love their own father.

- (11) a. Only Amina _{i} [t_i loves her father]
 b. Only Amina [λx [x loves y 's father]]
 c. Only Amina [λx [x loves x 's father]] *Binding*
 d. Only Amina [λx [x loves y 's father]] *Coreference* ($y = \underline{\text{Amina}}$)

On the coreference construal, the free variable y and the subject DP are assigned the same semantic value (11d). Since the individual Amina serves as a discourse referent for both the DP and the pronoun, the resulting VP denotes the set of female individuals who love Amina's father. The sentence thus asserts that only Amina belongs to this set – that is, other women do not love Amina's father. In sum, depending on how anaphora is resolved, different properties are predicated of the antecedent. Reinhart further assumes that when these two anaphora construals are truth conditionally equivalent, only binding survives.

D&UE make the case for this view of anaphora in the temporal domain, as stated in (12).

- (12) *Temporal anaphora*
 a. Anaphora between time-denoting arguments can be resolved via either (variable) binding or coreference
 b. Binding and coreference yield temporal construals that can be disambiguated in certain contexts.
 c. Binding is the default construal of anaphora

The assumption that aspects, just like tenses, order time intervals, also allows D&UE (2014) to extend the parallels established between tenses and pronouns to the domain of viewpoint **aspect**:

- (13) *Temporal anaphora and viewpoint aspect*
 a. Aspect establishes a viewpoint by ordering two times (AST-T, EV-T)
 b. This ordering relation can be established via anaphora.
 c. 'Coreference' yields a 'perfective' viewpoint, while 'binding' yields an 'imperfective' viewpoint.

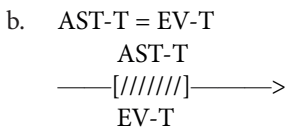
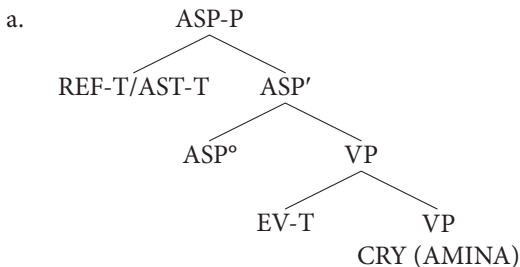
The two next subsections explicate the hypothesis in (13c).

2.3.1 Temporal coreference

D&UE thus adopt a referential approach to temporal phenomena involving pronominal like expressions referring to times. These temporal DPs (Zeit-Ps) are projected into the syntax as arguments of predicates T° , ASP° , and V° . Just like regular pronouns, temporal pronouns can be assigned a discourse referent (time interval) from the discourse storage. Coreference is the assignment of identical semantic values (discourse referents) to two individual denoting noun phrases (DPs). ‘Coreference’ in the temporal realm will likewise arise when two Zeit-Ps are assigned identical semantic/temporal values.

As explicated above, D&UE endorse the hypothesis that the ordering relations between temporal arguments can be established via anaphora. Suppose then that anaphora between the two arguments of ASP° (AST-T, EV-T) is resolved as coreference. What would the resulting viewpoint be? Perfective. To see why, consider (14).

(14) Temporal coreference

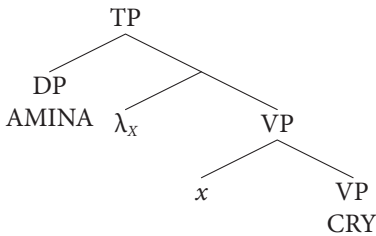


The Zeit-Ps projected in the specifiers of ASP° and VP° in (14a) are free temporal variables, which can be assigned a value from the discourse storage. If the AST-T and the EV-T are co-valued, they denote exactly the same temporal span. Co-valued times are thus coextensive (have the same initial and final bounds), since they are assigned exactly the same time interval as a semantic value. That is to say, when the temporal ordering between the arguments of ASP° is established via temporal anaphora and anaphora is coreference, the REF-T (the AST-T) ‘cannot denote a subinterval’ of the EV-T since they denote identical times: an imperfective viewpoint is thus automatically excluded, and the resulting viewpoint is what is called ‘perfective’ (‘aoristic’) aspect, the event is portrayed in its entirety, from its initial to its final bound.

2.3.2 Temporal binding

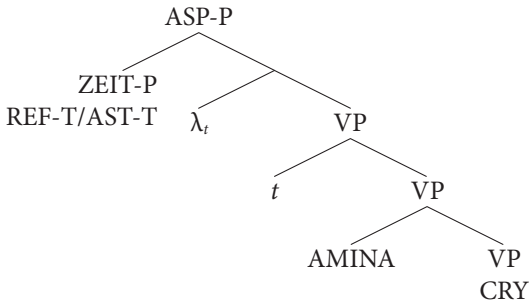
There is, however, an alternative for resolving anaphora: binding. Suppose then that anaphora between the two arguments of ASP° (AST-T, EV-T) is resolved as binding. The parallel between ordinary DP binding and Zeit-P binding is sketched below.

(15) a. DP Binding



a'. AMINA λ_x [CRY (x)]

b. Zeit-P/Temporal DP binding



b'. AST-T λ_{EV-T} [AMINA CRY (EV-T)]

λ -Binding of the variable in subject position in (15a) yields a predicate denoting the set of individuals of which the property of crying holds. This property is applied to the antecedent DP *Amina* in Spec TP. Parallely, binding of the time variable in the highest specifier of the VP (the EV-T) yields a predicate denoting the set of times of which the property of Amina's crying holds. This property is applied to the antecedent Zeit-P in Spec ASP-P (the AST-T).² Temporal binding thus ensures that the AST-T be a time at which Amina's crying holds. Activities like *cry* have the subinterval property (if Amina cried from 1:00 to 2:00 then most subintervals of this time are times at which Amina cried). Now, since (i) *cry* is an activity holding of all subintervals of the EV-T, and (ii) binding of the EV-T by the AST-T requires that the AST-T itself be a time at which *cry* holds, then as long as there is temporal overlap/coincidence between the AST-T and the EV-T, the AST-T will itself also

2. The reader should keep in mind that AST-T, EV-T, REF-T are merely convenient labels for time variables, used to help keep track of which time variable(s) are being talked about.

have the property of being an interval at which the activity of crying holds, as is the case in all of the temporal configurations illustrated in (16). The time for which an assertion is made ‘may’ (16c–d), but crucially ‘need not’ (16a–b) overlap with (include) the final bound of the event. Since the terminal bound of the event need not fall within the AST-T, binding yields an ‘imperfective’ viewpoint.

- (16) a.
$$\begin{array}{c} \text{AST-T} \\ -[-\text{//////////}-]-----> \\ \text{EV-T/crying} \end{array}$$
- b.
$$\begin{array}{c} \text{AST-T} \\ -\text{//////////}\{\text{//////////}-\}---> \\ \text{EV-T/crying} \end{array}$$
- c.
$$\begin{array}{c} \text{AST-T} \\ -\text{//////////}\{\text{//////////}\text{//////////}\}////-> \\ \text{EV-T /crying} \end{array}$$
- d.
$$\begin{array}{c} \text{AST-T} \\ ---[-\text{//////////}\}//////////-> \\ \text{EV-T/crying} \end{array}$$

Recapitulating, binding merely requires the crying time in (16) to overlap/coincide with the AST-T. This means that the AST-T can but need not exclude the bounds of the event. If the AST-T excludes the bounds of the event (16a), the event described is understood as ongoing at the reference-time. Since, however, either one or both bounds of the EV-T can fall within the AST-T (16b–d), they are ‘visible’, accessible to temporal modification. This is why modification by a completive (*in x-time*) adverbial in (9) above is felicitous, and why modification by a punctual adverb in (6a–b) yields (alongside the ongoing construal in (6i)) the sequential construal in (6ii), where the time denoted by the adverb specifies the (initial) bound of the described event, as is the case for activities/accomplishments under a perfective viewpoint (cf. (7)).

D&UE’s (2014) analysis of the imperfective vs. progressive nicely captures Molendijk’s (2005: 124–125) insight that the progressive characterizes an eventuality as ‘going on’, ‘in progress at the (past) REF-T’ (in D&UE’s terms, REF-T / AST-T within EV-T), thus excluding the bounds of the event (1b), while the *Imp* characterizes an eventuality that is ‘true/holds at a (past) REF-T’ (in D&UE’s terms, AST-T binds the EV-T), thus merely requiring that the AST-T/ REF-T be a time at which the EV-T be true/holds.

The next section extends D&UE’s analysis of imperfective vs. perfective past to Sequence of Tense (SoT) contexts, showing how this hypothesis accounts for one of the best known differences between imperfective and perfective aspect in subordinate contexts.

3. Imperfective vs. perfective past in embedded contexts

A notorious difference between imperfective and perfective is that only the former can be used to convey temporal simultaneity between a matrix and an embedded clause, as the contrast below illustrates. The *Imp* in (17) allows a simultaneous construal where the past time of laughing overlaps with the past saying-time, alongside a sequential construal where the time of laughing is past-shifted relative to the past matrix saying-time. In contrast, a simultaneous construal is unavailable when the embedded predicate is in the perfective: (18) only allows a sequential past-shifted construal.³

(17) a. (Spanish, Stowell 2014: 907)

Nora dijo que las chicas se reían de ella
 Nora said_{perf} that the girls REFL laughed_{imp} of her
 “Nora said that the girls were laughing at her.”

√Simultaneous √Sequential

b. (French)

Nora a dit que les filles riaient d'elle
 Nora has said that the girls laughed_{imp} of her
 “Nora said that the girls were laughing at her.”

√Simultaneous, √Sequential

(18) (Spanish, Stowell 2014: 907)

Nora dijo que las chicas se rieron de ella
 Nora said_{perf} that the girls REFL laughed_{perf} of her
 “Nora said that the girls laughed at her.”

*Simultaneous, √Sequential

Before tackling the question of how to explain this contrast, two opposing families of accounts of simultaneous construals are reviewed.

3.1 Sequence of Tense (SoT)

There are two major approaches to the simultaneous readings of a past tense in the complement clause of a past propositional attitude predicate, depending on whether the embedded tense is a semantic past tense, pragmatically competing with the present to express simultaneity (Gennari 2003; Altshuler & Schwarzschild 2013,

3. No French perfective counterpart to (18) is given because the simple perfective past in standard (spoken) French (the *passé simple*) is highly marked (mostly restricted to story-telling or narrative contexts). The strong tendency is to use the compound past (the *passé composé*, a present perfect) in lieu of the simple past, as shown with the contrast between Spanish (17a) with a simple perfective past (*preterito*) in the matrix clause, and French (17b) with a *passé composé* in the matrix in lieu of the simple past.

a.o.), or not interpreted as a semantic past (Kratzer 1998; Ogihara 1996; Stowell 2007, a.o.).

On the latter approach to SoT phenomena, the simultaneous and sequential past-shifted readings of past under past sentences arise via distinct LFs, as schematized in (19). On the past-shifted construal (19b), the morphological past in the embedded clause (just like the morphological past in the matrix) instantiates a true semantic past tense (henceforth, PAST) conveying anteriority of the embedded AST-T/EV-T relative to the matrix AST-T/EV-T. In contrast, on the simultaneous construal (19c), the morphological past in the embedded clause does not instantiate PAST since it does not induce past shifting relative to the time of the matrix AST-T/EV-T, but rather remains uninterpreted at LF. Different implementations of this proposal have been put forth: on Ogihara's (1996) tense deletion account, an embedded tense may be optionally deleted when it is c-commanded by a tense of the same kind, while on Kratzer's (1998) zero-tense account, the embedded tense is a temporal pronoun lacking temporal features of its own and, as such, interpreted via binding by its antecedent, the matrix past tense. On this class of proposals, there is thus a mismatch between the semantics of tense and its morphological expression under the simultaneous reading.⁴

- (19) a. *Cleo said that Kiya was sick*
 b. [Cleo PAST say [that Kiya PAST be sick]] → past-shifted
 c. [Cleo PAST say [that Kiya ~~PAST~~ /∅ be sick]] → simultaneous

In contrast, on pragmatic accounts of SoT phenomena, there is a one-to-one mapping between past tense morphology and meaning. That is, there is a single LF for (19a) involving a true PAST conveying anteriority in the complement clause, just like in the matrix. Semantically, the time of Kiya's being sick is thus past-shifted relative to the past time of Cleo's saying. However, when the embedded predicate is stative, the truth conditions of past-shifting will be compatible with a simultaneous construal because of the temporal profile of stative predicates: the time of Kiya's being sick can in principle extend forward to coincide with the matrix attitude time, thus yielding an apparent simultaneous construal.⁵

4. This class also includes accounts along the lines of Stowell (2007) which take past-under-past simultaneous readings as instantiating polarity or agreement phenomena. The idea is that a morphological past tense is subject to a licensing constraint requiring it to fall in the scope of a semantic PAST tense operator. A simultaneous construal arises because an embedded past can be long-distance licensed by a matrix (semantic) PAST.

5. 'Temporal Profile of Statives' (Altschuler & Schwarzschild 2013: 45):

For any tenseless stative clause φ and world w , if φ is true in w at moment m , then there is a moment m' preceding m at which φ is true in w and there is a moment m'' following m at which φ is true in w .

3.2 Why perfective (unlike imperfective) enforces past-shifting

Recall our hypothesis: when anaphora between the arguments of ASP° (AST-T, EV-T) is resolved as binding, the resulting viewpoint is imperfective, when it is resolved as coreference, the viewpoint is perfective. Let us turn to the question of why perfective past enforces past-shifted construals, unlike imperfective past which also allows simultaneous construals. Let's run through the temporal derivation of the Spanish perfective past in (20a), assigned the temporal syntax in (20e). In the matrix clause (20b), past orders the UT-T after the matrix AST-T (henceforth, AST-T1), itself coreferring with the matrix EV-T (EV-T1) since the viewpoint is perfective, yielding the temporal output in (20i). Turning to the complement clause (20c): TP2 gets temporally anchored into the matrix via binding of its highest REF-T (REF-T2, the external argument of the embedded tense) by the closest c-commanding time argument – that is, by EV-T1, the time of saying.⁶ Binding ensures that EV-T1 temporally coincides with REF-T2, as illustrated in (20ii). REF-T2 is then itself ordered after the embedded AST-T (AST-T2) by the embedded past (20iii). Finally, the ordering between AST-T2 and EV-T2 is established via coreference, yielding the temporal output in (20iv).

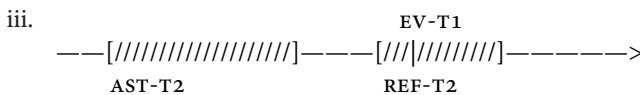
(20) a. *Cleo dijo que Kiya estuvo enferma*

Cleo said that Kiya was_{perf} sick

b. Matrix (TP1)



c. Complement (TP2)

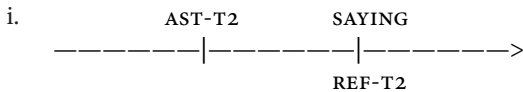


6. Unlike in independent/matrix clauses, where the external argument of T° denotes the UT-T (e.g., (3)), in complement clauses, the external argument of the subordinate tense is anaphorically dependent on the matrix event-time. Anaphora here is binding since the latter is the default construal of anaphora.

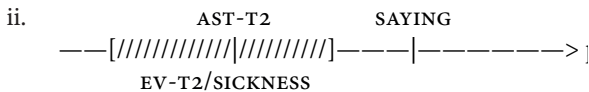
by AST-T2 requires that AST-2 be a time of which the state of Kiyā's being sick holds, as is the case in either of the temporal configurations in (21ii–iii). Since binding merely requires the time of Kiyā's being sick to overlap with AST-T2, the final bound of EV-T2 remains unordered relative to the final bound of AST-T2, and the time of Kiyā's being sick can (but need not (21ii)) extend forward to coincide with the matrix saying-time (EV-T1), thus allowing the simultaneous construal in (20iii).⁷

- (21) a. *Cleo dijo que Kiyā estaba enferma*
Cleo a dit que Kiyā était malade
 Cleo said that Kiyā was_{imp} sick

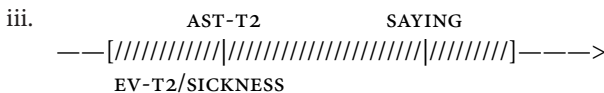
b. Output of (20i)–(20iii)



c. Imperfective viewpoint



✓past-shifted



✓simultaneous

The proposal that anaphora between the AST-T and the EV-T (the time arguments that ASP^o orders) can be resolved as either coreference (yielding a perfective viewpoint) or binding (yielding an imperfective viewpoint) straightforwardly explains why perfective past under past embeddings enforce past-shifted readings, while imperfective past under past embeddings allow both simultaneous and past-shifted construals.

Note, however, that with imperfective past under past embeddings, the simultaneous reading is the default construal (the construal arising spontaneously in out of the blue contexts). But why is the simultaneous reading preferred over the past-shifted reading? Now, the very same question arises on a zero-tense/tense deletion accounts of SoT. On the proposal defended here, there are two different tense/aspect forms that can be used to express the past-shifted reading (imperfective and perfective past), but only one that can be used to express the simultaneous reading (imperfective past). This suggests that the default status of the simultaneous

7. This account generalises to predicates with the subinterval property, as explained in § 2.3.2 in the discussion of activities such as *cry*.

construal with the imperfective past can be accounted for on the assumption that perfective and imperfective past compete to express the past-shifted reading. In a nutshell, the use of a statement with an embedded imperfective past, when the corresponding more informative, stronger statement with an embedded perfective past could have been used, triggers the inference that the stronger statement is false – that is, that the past-shifted reading does not hold.

4. Strict vs. sloppy temporal construals

Having extended D&UE's (2014) proposal for perfective vs. imperfective past to SoT contexts, we are now in a position to test this proposal by appealing to a standard diagnostic for distinguishing coreference anaphora from BVA: ellipsis. The idea of appealing to temporal TP ellipsis as a diagnostic tool is inspired by and builds on Stowell's (2014) commentary on D&UE (2014). Ellipsis has been used to elucidate the interpretation of pronouns, on the classic assumption that coreference yields strict identity readings under ellipsis, while BVA yields sloppy identity readings (Sag 1976). Stowell probes the referential behavior of imperfective vs. perfective past in subordinate contexts to see whether it reveals a distinction between strict and sloppy identity interpretations, as would be expected if indeed Reinhart's distinction between coreferential and bound variable interpretations is applicable to time-denoting arguments, as D&UE contend.⁸

Before proceeding, Reinhart's view of anaphora should be explicated further since she makes two levels of distinction in how anaphora can be resolved: (i) binding vs. coreference, and (ii) coreference with a referential DP vs. with a bound variable. Coreference is thus available regardless of the referential status of the antecedent, as the interpretations of the seminal example from Heim (1998) and Reinhart (1997) given in (22) demonstrate. Anaphora in (22a) is BVA since the antecedent of both pronouns in the embedded clause is a QP. (22a) is nonetheless ambiguous since two different properties can be ascribed to the embedded subject pronoun, just as they could in (22b) (or (8) above) where the antecedent of the pronoun is a referential DP. To see this, consider first what property must hold of every girl for (22a) to be true under the binding construal in (23a), where the embedded possessive pronoun is a variable locally bound by the embedded subject/ λ -operator (cf. Reinhart's definition of logical-syntax binding in footnote 1). The property in question is that of loving one's own father. On this construal, (22a) entails that

8. Stowell (2014) does not actually test the predictions that D&UE (2014) make for ellipsis, for the simple reason that D&UE do not discuss (past-shifted vs. simultaneous readings in) SOT contexts in the first place.

every girl thinks that I, for instance, do not love my own father. Consider next what property must hold of every girl for (22a) to be true under the coreferential construal in (23c), where the embedded possessive is a variable co-valued with the embedded subject (itself a variable bound by the matrix λ -operator). Assuming Amina is one of the girls in the context, the property in question is that of loving Amina's father. (22a), under this construal, entails that every girl thinks that I, for instance, do not love Amina's father. In sum, binding and coreference do not have the same truth conditions, whether the matrix subject is a QP as in (22a), or a referential DP as in (22b).

- (22) a. *Every girl₁ thinks that only she₁ loves her₁ father*
 b. *Amina₁ thinks that only she₁ loves her₁ father*

(23) *Binding:*

- a. *Every girl ($\lambda x(x$ thinks only $x(\lambda y(y$ loves y 's father)))) \rightarrow binding of *her**
 b. *Amina ($\lambda x(x$ thinks only $x(\lambda y(y$ loves y 's father)))) \rightarrow binding of *her**

Coreference:

- c. *Every girl ($\lambda x(x$ thinks only $x(\lambda y(y$ loves x 's father)))) \rightarrow *her* covalued with a bound variable*
 d. *Amina ($\lambda x(x$ thinks only $x(\lambda y(y$ loves x 's father)))) \rightarrow *her* covalued with a referential DP*

The null assumption is that this three-way contrast in how anaphora can be established – that is, via A-binding as in (23a–b) vs. covaluation with a bound variable as in (23c), or a referential DP as in (23d) – extends to temporal anaphora. This is indeed the case, as I shall now show by probing the distribution of temporal sloppy vs. strict readings in SoT contexts. First, a few words on how to account for temporal ellipsis.

4.1 Temporal ellipsis: Structural vs. referential parallelism

I adopt the classic assumption that ellipsis is subject to an 'LF-parallelism constraint' (e.g., Fox 2000), with strict readings arising via referential parallelism, and sloppy ones via structural parallelism, moreover extending this assumption to temporal ellipsis. The effect of parallelism is illustrated with (24). Parallelism requires the pronoun *her* in both the antecedent and elided VPs either to be assigned the same referential value ('Referential Parallelism'), yielding the strict reading in (24a), or to be linked to their (respective) antecedent via identical dependencies ('Structural Parallelism'), yielding the sloppy reading in (24b). My specific proposal for extending the parallelism constraint from the nominal to the temporal domain is given in (25), on the default hypothesis in (26) that with time denoting arguments, just like

with individual denoting arguments, strict readings arise via referential parallelism, while sloppy readings arise via structural parallelism.

- (24) *Cleo₁ loves her₁ mother and Kiya does too*
- Referential parallelism* → *strict reading*
Cleo₁ loves her₁ mother and Kiya ~~loves her₁ mother~~ too
 - Structural parallelism* → *sloppy reading*
Cleo₁ loves her₁ mother and Kiya₂ ~~loves her₂ mother~~ too
- (25) *Parallelism Constraint on Temporal ellipsis*
The EV-Ts in the antecedent and elided constituents must either
- have the same referential value (Referential Parallelism), or
 - show matching ordering relations (Structural Parallelism)
- (26) Referential parallelism yields strict readings, while structural parallelism yields sloppy readings, be it in the nominal or the temporal domain.

I now show how (25)–(26) together with the analysis of past-shifted vs. simultaneous construals in SoT contexts put forth in § 3 account for the distribution of temporal sloppy vs. strict readings of (im)perfective past in SoT contexts.

4.2 Perfective past in ellipsis contexts

Consider the coordinate structure in (27) adapted from Stowell (2014), where the second conjunct contains an elided VP/TP.⁹ The first conjunct, which serves as antecedent for the elided TP, contains a perfective past (in the complement clause), itself embedded under a matrix (perfective) past. As Stowell points out, (27) only allows the past-shifted construals in (27a–b), where the time of sickness precedes the time of saying, be it in the antecedent or the elided clause, as diagrammed in (28a).

- (27) *Cleo le dijo a su jefe que estuvo enferma y Kiya*
Cleo him told to her boss that was_{perf} sick and Kiya
(*debía*) ___ *también*
(*must_{imp}*) ___ *too*
“Cleo told her boss that she was sick and Kiya (should have) too.”
- Strict past-shifted reading*
Cleo told her boss at t_1 that she was sick at t_2
[sick₂ – saying₁ – UT-T]
& Kiya told her boss at t_3 that she was sick at t_2
[sick₂ – saying₃ – UT-T]

9. As Stowell (2014) himself points out, ellipsis in (27) may involve TP rather than VP ellipsis per se, since Spanish (and likewise French) lacks VP ellipsis of the English variety and modal ellipsis thus has to be used as a diagnostic tool instead.

b. *Sloppy past-shifted reading*Cleo told her boss at t_1 that she was sick at t_2 [sick₂ – saying₁ – UT-T]& Kiya told her boss at t_3 that she was sick at t_4 [sick₃ – saying₄ – UT-T](28) a. *Perfective past under past*

AST-T2=EV-T2	SAYING	UT-T	
— [//////////]	—	—	—>past-shifted (only)

b. *Referential parallelism* → *strict* past-shifted reading (27a)c. *Structural parallelism* → *sloppy* past-shifted reading (27b)

Now, this state of affairs is exactly as expected since, as established in § 3.2, with a perfective viewpoint, the simultaneous construal is automatically excluded (EV-T2 in (28a) cannot extend forward, beyond the final bound of AST-T2, to coincide with the matrix saying-time). By hypothesis, temporal ellipsis is subject to a parallelism constraint that can be satisfied in either of two ways. If this constraint is satisfied via referential parallelism – that is, if EV-T2 in both the antecedent and elided VPs are assigned the same referential value – then the *strict* past-shifted reading in (27a), where Cleo and Kiya are sick at the very same past-shifted time (t_2), arises. If this constraint is satisfied via structural parallelism – that is, if the ordering relations of EV-T2 in the antecedent and elided clause (relative to their respective higher REF-T) match – then the *sloppy* past-shifted reading in (27), where Cleo and Kiya are sick at different past-shifted times (t_2 and t_4 respectively), arises.

4.3 Imperfective past in ellipsis contexts

Let's now turn to Example (29), where the antecedent of the elided clause this time contains an embedded imperfective past, and which, unlike its perfective counterpart in (27), yields a variety of past-shifted and simultaneous readings:

- (29) *Cleo le dijo a su jefe que estaba enferma y Kiya*
Cléo a dit à son patron qu'elle était malade et Kiya
 Cleo him told to her boss that-she was_{imp} sick and Kiya
 (debía) — también
 (devait) — aussi
 (must_{imp}) too

“Cleo told her boss that she was sick and Kiya (should have) too.”

a. Cleo told her boss at $t1$ that she was sick at $t1$ & Kiya told her boss at $t2$ that she was sick at $t2$ √sloppy simultaneous

- b. ✓ *Sloppy past-shifted* (29b):¹⁰
 Cleo and Kiya are sick at different past times
 Time of sickness past-shifted wrt saying-time in both antecedent & elided TPs
- c. **Sloppy mixed* (29c):
 Cleo and Kiya are sick at different past times → No referential parallelism
 Time of sickness overlaps the saying-time in the antecedent TP, but past-shifted in the elided TP → No structural parallelism

I now turn to the question of why (29) can also yield the range of strict readings in (29d–f). By hypothesis, strict readings arise under ‘referential’ parallelism: when EV-T2 in the elided TP is assigned the same temporal value as EV-T2 in the antecedent TP. Now, since the viewpoint is imperfective, AST-T2 binds EV-T2 in the antecedent clause. This means that EV-T2 in the antecedent clause is itself a bound variable. In other words, co-valuing EV-T2 in the elided TP with EV-T2 in the antecedent TP instantiates covaluation (that is, coreference) with a BV. This option yields the range of strict readings in (32).

- (32) *Referential* parallelism → strict temporal readings
- a. ✓ *Strict* past-shifted ((29d)):
 Cleo and Kiya are sick at the same past time
 Time of sickness past-shifted wrt saying-time in both antecedent & elided TP
- b. ✓ *Strict* simultaneous ((29e)):
 Cleo and Kiya are sick at the same past time
 Time of sickness overlaps saying-time in both antecedent & elided TP
- c. ✓ *Strict* mixed ((29f))
 Cleo and Kiya are sick at the same past time → ✓Referential parallelism
 Time of sickness overlaps saying-time in the antecedent TP, but past-shifted in the elided TP → No structural parallelism

Under all of these readings, Cleo and Kiya are sick at the same past time, thus satisfying referential parallelism. The ‘strict past-shifted’ reading (29d)/(32a) arises via referential parallelism when EV-T2 is past-shifted relative to EV-T1 (as diagrammed in (30b)) in both the antecedent and elided clauses. The ‘strict simultaneous’ reading (29e)/(32b) arises via referential parallelism when EV-T2 is simultaneous to EV-T1 (as diagrammed in (30a)) in both the antecedent and elided clauses. Finally, the

10. Recall that the simultaneous reading is the default reading for an embedded imperfective past, but the sloppy past-shifted reading becomes available with sufficient contextual information – e.g., *there was a meeting this morning, all the employees know that Cleo and Kiya were sick last week*. Later on, A asks B what happened at the meeting and B responds with (29).

proposal also allows for the reading identified by Stowell (2014) in (29f)/(32c), labeled here ‘*strict mixed*’, where the embedded state-time is understood as simultaneous with the matrix saying-time in the overt antecedent clause, but as past-shifted in the elided clause. This mismatch is allowed because it satisfies parallelism referentially. Recall, crucially, however, that the converse ‘*sloppy mixed*’ reading in (29c)/(31c) was unavailable. This followed automatically on the account put forth here since, in that case, parallelism could be satisfied neither structurally (the ordering relations in the antecedent and elided clauses do not match), nor referentially (the times at which Cleo and Kiya are sick are disjoint).

5. Conclusion

The framework developed by D&UE incorporates insights from both referential and predicative theories of tense. Under this mixed approach, temporal coincidence between time arguments can be established in a variety of ways, each yielding a distinct temporal viewpoint (§ 2). In line with predicative theories, temporal coincidence can be established via a predicate of spatiotemporal ordering expressing inclusion between the time arguments it relates. When this predicate is generated under ASP^o, it orders the AST-T ‘within’ the EV-T, yielding a progressive viewpoint. In line with referential theories, temporal coincidence can also arise via temporal anaphora. If anaphora between the arguments of ASP^o is resolved as binding, the viewpoint is imperfective (binding ensures that the AST-T and the EV-T temporally overlap). Alternatively, if anaphora between the arguments of ASP^o is resolved as coreference, the viewpoint is perfective (coreference ensures that the AST-T and EV-T denote exactly the same time interval).

Novel evidence was adduced for D&UE’s approach to (im)perfective aspect, by showing how it accounts for the construals of perfective vs. imperfective past in SoT contexts (§ 3), as well as predicting the intricate distribution of strict vs. sloppy readings in SoT contexts under ellipsis (§ 4), recapitulated in (33):

(33) *Temporal Ellipsis Generalisations (TEG)*

Under ellipsis,

a. Perfective past yields:

strict or *sloppy* past-shifted, but **not** *simultaneous*, readings

b. While imperfective past yields:

strict or *sloppy* simultaneous, as well as past-shifted, readings

strict but **not** *sloppy mixed* readings

The TEG is predicted by the interplay of independently motivated assumptions. A perfective viewpoint arises via coreference, which precludes a simultaneous

construal in SoT contexts, enforcing a past-shifted interpretation (§ 3). It is thus expected that, under ellipsis, perfective past will only allow past-shifted readings, while imperfective past also allows simultaneous and mixed readings. The distribution of strict vs. sloppy readings follows from the proposal defended in § 4 to extend the parallelism constraint on nominal ellipsis to temporal ellipsis, adapting to the temporal domain the classic assumption that ‘referential’ parallelism yields ‘strict’ readings, while ‘structural’ parallelism yields ‘sloppy’ readings. Finally, Reinhart’s (1997) contention that coreference is available irrespective of the referential status of the antecedent extends to temporal anaphora since strict readings with an imperfective past involve covaluation with a bound EV-T, while strict readings with a perfective past involve covaluation with a referential EV-T.

Acknowledgements

This chapter is based on D&UE (2014) and Stowell’s (2014) comments on the former. I thank the organizers of the GREAT workshop for their invitation to Utrecht, which always brings back fond recollections of stimulating meetings with Tanya Reinhart. I take this opportunity to dedicate this paper to her memory, as it seeks specifically to extend Tanya’s view of anaphora to the temporal domain and, in particular, to viewpoint aspect. Thanks also to Tim Stowell and Myriam Uribe-Etxebarria for inspiring discussion. All shortcomings and errors are my own.

References

- Altshuler, Daniel. 2016. *Events, states and times: An essay on narrative discourse in English*. Berlin: de Gruyter. <https://doi.org/10.1515/9783110485912>
- Altshuler, Daniel & Roger Schwarzschild. 2012. Moment of change, cessation implicatures and simultaneous readings. In Emmanuel Chemla, Vincent Homer & Grégoire Winterstein (eds.), *Proceedings of Sinn und Bedeutung* 17, 45–62.
- Comrie, Bernard. 1976. *Tense*. Cambridge: Cambridge University Press.
- Demirdache, Hamida & Myriam Uribe-Etxebarria. 2007. The syntax of time arguments. *Lingua* 117. 330–366. <https://doi.org/10.1016/j.lingua.2005.07.006>
- Demirdache, Hamida & Myriam Uribe-Etxebarria. 2014. Aspect and temporal anaphora. *Natural Language and Linguistic Theory* 32. 855–895. <https://doi.org/10.1007/s11049-014-9231-2>
- Enç, Murvet. 1987. Anchoring conditions for tense. *Linguistic Inquiry* 18. 633–657.
- Fox, Danny. 2000. *Economy and semantic interpretation*. Cambridge, MA: MIT Press.
- Gennari, Sylvia. 2003. Tense meaning and temporal interpretation. *Journal of Semantics* 20. 35–71. <https://doi.org/10.1093/jos/20.1.35>
- Heim, Irene. 1998. Anaphora and semantic interpretation: A reinterpretation of Reinhart’s approach. In Orin Percus & Uli Sauerland (eds.), *The interpretive tract* (MIT Working Paper in Linguistics 25), 205–246. Cambridge, MA: MITWPL.

- Kamp, Hans & Christian Rohrer. 1983. Tense in texts. In Rainer Bäuerle, Christoph Schwartze & Armin von Stechow (eds.), *Meaning, use and interpretation of language*, 250–269. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110852820.250>
- Kauf, Carina & Hedde Zeijlstra. 2018. Towards a new explanation of sequence of tense. *Proceedings of SALT* 28, 59–77. <https://doi.org/10.3765/salt.v28i0.4407>
- Klein, Wolfgang. 1994. *Time in language*. London: Routledge.
- Klein, Wolfgang. 2000. An analysis of the German perfekt. *Language* 76(2), 358–382. <https://doi.org/10.1353/lan.2000.0140>
- Kratzer, Angelica. 1998. More structural analogies between pronouns and tenses. *Proceedings of SALT* 8, 92–110. <https://doi.org/10.3765/salt.v8i0.2808>
- Laca, Brenda. 2005. Périphrases aspectuelles et temps grammaticaux dans les langues romanes. In Hava Bat-Zeev Shyldkrot & Nicole Le Querler (eds.), *Les périphrases verbales* (Linguisticae Investigationes Supplementa 25), 47–66. Amsterdam: John Benjamins. <https://doi.org/10.1075/lis.25.06lac>
- Molendijk, Arie. 2005. The imparfait of French and the past progressive of English. In Bart Hollebrandse, Angeliek van Hout, & Co Vet (eds.), *Crosslinguistic views on tense, aspect and modality* (Cahier Chronos 13), 119–129, Amsterdam: Rodopi. <https://doi.org/10.1017/S0959269507002918>
- Ogihara, Toshiyuki. 1996. *Tense, attitudes and scope*. Dordrecht: Kluwer. <https://doi.org/10.1007/978-94-015-8609-2>
- Partee, Barbara. 1973. Some structural analogies between tenses and pronouns in English. *The Journal of Philosophy* 70, 601–609. <https://doi.org/10.2307/2025024>
- Reinhart, Tanya. 1997. *Strategies of anaphora resolution*. Uil OTS Working Papers, TL97-007. Utrecht: Utrecht University.
- Sag, Ivan. 1976. Deletion and logical form. Ph.D. Dissertation: Massachusetts Institute of Technology.
- Schaden, Gerhard. 2009. *Composés et surcomposés. Le “parfait” en Français, Allemand, Anglais et Espagnol*. Paris: L’Harmattan.
- Smith, Carlota. 1991. *The parameter of aspect*. Dordrecht: Kluwer. <https://doi.org/10.1007/978-94-015-7911-7>
- Stowell, Tim. 2007. The syntactic expression of tense. *Lingua* 117, 437–463. <https://doi.org/10.1016/j.lingua.2005.08.003>
- Stowell, Tim. 2014. Capturing simultaneity: A commentary on the paper by Hamida Demirdache & Myriam Uribe-Etxebarria. *Natural Language and Linguistic Theory* 32, 897–915. <https://doi.org/10.1007/s11049-014-9241-0>

The structure and interpretation of ‘non-matching’ split interrogatives in Spanish

Olga Fernández-Soriano
Universidad Autónoma de Madrid

The goal of this paper is to analyze the properties of (a special type of) ‘split interrogative’ (SI) constructions in Spanish. SIs are *wh*-questions followed by a phrase that constitutes a possible answer, the ‘tag’. The overall structure is interpreted as a yes/no question (as in *what did John bring, a book?*). In standard cases, the tag matches the (case and thematic) features of the *wh*-element. Nevertheless, in (spoken Peninsular) Spanish what I will call ‘Non-matching Split Interrogatives’ (NMSI) are also possible. In these cases, the *wh*-element and the XP in the tag may not match; instead, it is the dummy (neuter) *qué* ‘what’ that heads the *wh*-clause. I investigate these cases and propose a (biclausal) analysis involving an ellipsis process similar to the one taking place in fragments (Merchant 2004). To support this hypothesis, I focus on the fact that: in NMSI there is a form-meaning mismatch that, to my knowledge, has gone unnoticed both in theoretical and descriptive studies.

Keywords: split interrogatives, focus, Spanish, neuter *qué*, ellipsis, fragment answers, scope marker

1. Split interrogatives

Split interrogatives contain a *wh*-question part and a tag that, in a still intuitive sense, provides a possible answer to that *wh*-question. The whole structure is interpreted as a yes/no question, which seeks to confirm the identity between the *wh*-operator and the tag. I provide some Spanish examples in (1b–e):

- (1) a. *¿Qué quiere el niño, un helado?*
“What does the kid want, an ice cream?”
- b. *¿Dónde lo pusiste, en la mesa?*
“Where did you put it, on the table?”
- c. *¿Adónde te fuiste, a casa?*
“Where did you go, home?”

- d. *¿Cómo/con qué lo has resuelto, con la calculadora?*
 “How/with what did you solve it, with a calculator?”
- e. *¿A quién se lo dijiste, a tu jefe?*
 “To whom did you tell (about it), to your boss?”

In the examples in (1), the tag is a correlate of the *wh*-phrase, with same case and thematic role. I will call these sentences ‘Matching Split Interrogatives’ (MIS). There are essentially two main alternative approaches for MIS: on the one hand, authors like Camacho (2002) propose a monoclausal structure where the tag and the *wh*-element form a constituent. The representation for (1a) would be like (2), adapted from (Camacho 2002: 162), where *qué* is an interrogative operator.¹

- (2) [C *qué*_i quiere el niño [_{DP1} [_{DP1} *h*_i] [_{DP} un helado]]]

On the other hand, authors like Arregi (2010) propose a biclausal analysis whence the two parts of the MSI are independently generated. The tag is the result of ellipsis in an interrogative (a yes/no question) with focus fronting to C, followed by deletion of the TP complement at PF. The licensing antecedent for ellipsis is the *wh*-element. I take the representation from Arregi (2010: 542):

- (3) $wh\text{-phrase}_i C_{Q,wh} \dots t_i] [CP_2 \text{tag}_j C_Q [TP \dots t_j \dots]]$
-

One of the reasons to propose a biclausal analysis has to do with prosody. The two parts of a split interrogative have the intonation contour of two separate questions: the *wh*-part has an initial pitch rise associated to the *wh*-word, followed by a gradual descent and ends with a sentence-final fall, whereas the tag has the intonation contour of a non-*wh*-question (final high tone).² In my view, the most important argument to support the biclausal analysis is that the non-focal part does not have to be elided:

- (4) a. *¿Qué quiere el niño, un helado quiere el niño?*
 “What does the child want, an ice cream the child wants?”
- b. *¿Qué le vas a decir, que no puedes ni verlo le vas a decir?*
 “What are you going to tell him, that you cannot stand him you are going to tell him?”

1. It is not clear if Camachos’s approach applies to MSI like (1) or could also extend to non-matching constructions studied in this work, since all the examples provided involve direct objects that, as will be seen, take the *wh*-form *qué* and no indication of intonation or interpretation is provided.

2. Nevertheless, there seems to be dialectal variation with respect to the intonation pattern of these sentences (López Cortina 2009).

Another reason for the biclausal analysis for the structures in (1) is that there are connectivity effects between the tag and the *wh*-clause. Of special relevance to us is the fact that ‘Case-matching-effects’ are obtained: in (1) the tags and the *wh*-elements have the same case marking (and are introduced by the same P if they are PPs).

The focus of this paper is that, together with matching structures as those in (1), in (Peninsular) Spanish split interrogatives may show a mismatch between the *wh*-element and the phrase in the tag. This fact has received little attention in theoretical and descriptive studies, although it is completely productive. I call these constructions ‘Non Matching Split Interrogatives’ (NMSI) and they are described in the next section.

2. Non matching split interrogatives

2.1 Properties

NMSI are not headed by a complex *wh*-element, corresponding to the focused phrase in the tag, but by the dummy form *qué* (“what”). Some examples, taken from real oral speech, are provided in (5):³

- (5) (Spont., Madrid)
- a. *¿Qué lo pusiste, en la mesa y se rompió?* (cf. (1b))
lit. “What did you put it, on the table (and it broke)?”
 - b. *¿Qué lo has resuelto, con una calculadora?* (cf. (1c))
lit. “What did you solve it, with a calculator?”
 - c. *¿Qué te fuiste, a casa?* (cf. (1d))
lit. “What did you go, home?”
 - d. *¿Qué se lo dijiste, a tu jefe?* (cf. (1e))
lit. “What did you tell about it, to your boss?”
 - e. *¿Qué se lo explicaste, después de clase?*
lit. “What did you explain it to her, after class?”

A property shared by both NMSI and MSI is that they cannot be embedded: they are both root phenomena, involving the complete left periphery. They do not satisfy selectional properties of interrogative verbs:

- (6) a. *Te preguntó adónde ibas / si ibas a casa.*
“He asked you where you were going / if you were going home”

3. All sentences in (5), and those marked as ‘Spont.’ throughout the paper, come from oral spontaneous Madrid speech and have been collected by the author.

- b. *Te preguntó a dónde ibas, a casa.
 You he-asked where you.were.going home
- c. *Te preguntó qué ibas, a casa.
 You he-asked what you.were.going, home

In addition, both in NMIS and MSI the *wh*-element can be preceded by left dislocated material, as seen in (7). It thus seems that both the complex *wh*-element in MSI and the neuter *qué* in NMSI occupy the position where *wh*-elements appear in the left periphery (SpecFocP).

- (7) a. *Juan ¿qué/adónde va, a casa?*
 Juan what/where he-goes, home
- b. *Tú ¿qué/cómo lo resuelves, con calculadora?*
 You, what/how you-solve it with a calculator
- c. *Tú ¿qué eres, el jefe?* (Spont., Madrid)
 You what are.you the boss
 “What about you, are you the boss?”

Despite these similarities, Arregi (2010) suggests that NMSI have a different analysis. I will show that, in fact, there are both phonetic (intonational) and interpretive differences between NMSI and MSI as well as other crucial properties that call for two different approaches.⁴

As for intonation, López Cortina (2009) describes a rising final tone in the first part of the interrogative in MSI as opposed to a lowering one for NMSI. He also claims that while a NMSI “is about its second part, with the first part presenting a context assumed to be true”, in MSI there are two juxtaposed questions, and the *wh*-clause “poses a question and then narrows its range by adding a second, more specific one” (López Cortina 2009: 221). Irurtzun (2017), who studies similar cases in Basque, shows that NMSI have a particular reading, which the author dubs ‘evidential’, in the sense that it has a confirmational interpretation (for which the speaker has direct or inferential evidence). In what follows, I describe the different approaches to this phenomenon.

2.2 The structure of NMSI. Monoclausal analyses

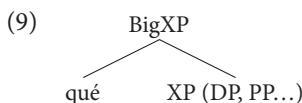
In general, the (few) authors who have analyzed NMSI coincide in defending a monoclausal structure. Contreras & Roca (2007) (C&R) analyze NMSI in Catalan and argue that the non-matching *wh*-element and the focus form a kind of ‘BIG

4. An anonymous reviewer asks if it is possible to distinguish between MSI and NMSI in cases where the *wh*-word is *qué*. The answer to this question is precisely that the intonation and (crucially) the interpretation are different.

DP’, similar to the one proposed for clitic doubling (Uriagereka 1995). The fact that *qué* is analyzed as a clitic explains, according to the authors, an important property of NMSI, namely, that nothing (including negation) can intervene between the *wh*-element and the verb. The same is true for Spanish:⁵

- (8) a. *¿Qué, según Pedro, iremos, al cine?
lit. “What, according to Peter, shall we go, to the movies?”
 b. *¿Qué no lo hiciste, con las manos?
lit. “What not didn’t you do it, with your hands?”

The mentioned authors propose a representation like (9), where a BigXP contains both the interrogative particle *qué* and the clefted element, the tag:



The phrase in the right node moves to the FocP Specifier, while *qué* cliticizes to IP and checks *wh*-features in ForceP. For a Catalan sentence like (10a) C&R propose a structure like (10b) and a complex derivation, with various steps, sketched in (10c):

- (10) a. *Què anirem, al cine?*
lit. “What shall we go, to the movies?”
 b. [IP I [anirem [_{BigXP} [qué] [al cine]]]]
 c. [_{ForceP} què_i Force⁰ [_{GroundP} [IP t_i [VP anirem [_{BigXP} t_i t_j]]]]_k Ground⁰[_{FocP} [al cine]_j Foc⁰ t_k]]]

C&R propose a ‘Ground Phrase’ between ForceP and FocP, following Munaro & Pollock (2005), who state that this phrase contains given information in interrogative sentences. I will not go into technical details of this complex derivation. The central idea is that *qué* is generated in a doubling construction, inside a monoclausal structure, and is further moved to Spec, ForceP, in the left periphery. GroundP accounts for the required interpretation.

Lorenzo (1994–95) also proposes a monoclausal analysis of NMSI. In this case, *qué* is considered an expletive, merged at the specifier of the clause. This expletive determines the scope of the tag, which appears as an adjunct.

Irurtzun (2017), in a (monoclausal) analysis for Basque NMSI proposes a predicative small clause (a Relator Phrase, in the sense of den Dikken (2006)) containing a *wh*-variable with default morphology, *zer* (“what”), and a tag containing the “answer”, which behaves as the focus. I reproduce below Irurtzun’s structure:

5. In this respect, they do not seem to contrast with MSI. The equivalent structures with matching *dónde* “where” and *cómo* “how” are equally out.

- (11) [
- _{RelP}
- Wh [
- _{Rel'}
- Rel Answer]]

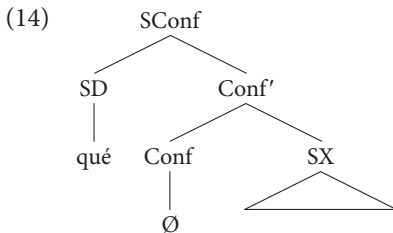
An important issue is that this predicative structure gives rise to a ‘confirmational interpretation’, since the *wh*-variable coincides with the potential focus. In particular, a relation is established between the interpretation of NMSI and that of clefts. Irurtzun provides the following example, with the corresponding interpretation.

- (12) *Zer zator, bihar?*
 what you.come tomorrow
 “Is it tomorrow that you are coming?”

In the Spanish equivalent constructions this interpretation is also obtained. In contrast with (13a), which can be paraphrased as in (13a'), (13b) is interpreted as in (13b').

- (13) a. *¿Cómo lo resolviste, con calculadora?*
 “How did you solve it, with a calculator?”
 a'. *¿Cómo lo resolviste, lo resolviste con calculadora?*
 “How did you solve it, did you solve it with a calculator?”
 b. *¿Qué lo resolviste, con calculadora?*
lit. “What did you solve it, with a calculator?”
 b'. *¿Fue con calculadora como lo resolviste, (verdad)?*
 “Was it with a calculator that you solved it, (right)?”

Finally López Cortina (2009) also proposes a monoclausal analysis, in this case with a ‘doubling’ structure. This author claims that *qué* in NMSI is generated in the specifier of an Extended Projection (Grimshaw 2005), which he calls ‘Confirmation Phrase’.⁶ This projection introduces the notion of ‘need for confirmation’ associated to a [+wh] feature which causes *wh*-movement and accounts for the particular interpretation obtained in NMSI. Confirmation Phrase (headed by a null ‘confirmation marker’) dominates the node that ends up being the ‘proposed answer’, the tag:



(López Cortina 2009: 224)

6. According to LC, this fact explains the adjunct-like behavior of *qué*, which seems to be sensitive to strong islands.

I will adopt some aspects of the preceding approaches. Nevertheless, as it will be seen in the next section, there are arguments that support a biclausal analysis, with ellipsis, for NMSI (parallel to the one proposed by Arregi (2010) for MSI). The most important one has to do with interpretation.

3. A biclausal analysis for NMSI. Ellipsis and ‘extended’ reading

3.1 Properties of NMSI

A very important property that distinguishes NMSI from MSI is that, in the former, the focused constituent (the tag) is always obligatory:

- (15) a. *¿Qué vas? (vs. ¿Dónde vas?)
lit. “What are you going? (vs. where are you going?)”
 b. *¿Qué lo has resuelto? (vs. ¿Cómo lo has resuelto?)
lit. “What did you solve it? (vs. how did you solve it?)”
 c. *¿Qué lo pusiste? (vs. ¿Dónde lo pusiste?)
lit. “What did you put it? (vs. where did you put it?)”
 d. *¿Qué se lo dijiste? (vs. ¿A quién se lo dijiste?)
lit. “What did you tell about it? (vs. whom did you tell about it?)”

What this contrast suggests is that, even if NMSI seem to involve two clauses, there must be a (selectional) relation between them (i.e., they are not independent). Other differences between NMSI and MSI indicate that the former deserve a different analysis. In particular, it seems that NMSI involve a special projection, related to discourse (be it Predicative, Speech Act or other) containing the focalized element. At the same time, there are reasons to believe that (as in MSI) the tag may initially be clausal and that an ellipsis process of all non-focal material has applied. I try to combine these two apparently contradictory facts in what follows.

The first argument that supports the biclausal status of NMSI is that, as was the case for MSI (see above), they allow the complete clause to appear in the tag, that is, the tag does not have to be restricted to the focused XP:

- (16) a. ¿Qué vas a ir ahora, a casa a dormir vas a ir?
lit. “What are you going now, home to sleep are you going?”
 b. ¿Qué lo resolviste, con una calculadora lo resolviste?
lit. “What did you solved it, with a calculator you solved it?”
 c. ¿Qué se lo vas a contar, al jefe se lo vas a contar?
lit. “What are you going to tell about it, the boss are you going to tell about it?”

Another important piece of evidence is that, together with argument and adjunct XPs, the tag in split interrogatives can contain the predicate or the whole sentence. In these cases, the main *wh*-sentence contains a proform with the verb *hacer* ‘to do’, or *pasar* ‘to happen’. Crucially, in these cases the *wh*-element heading the interrogative is always the neuter form *qué*. See the following standard Spanish examples:

- (17) a. *¿Qué hiciste, irte a tu casa?*
 “What did you do, go home?”
 b. *¿Qué pasó, que se puso a llover?*
 “What happened, it started to rain?”
 c. *¿Qué pasa, (que) te sientes incómoda?*
 “What happens/What’s wrong, you are feeling uncomfortable?”

My analysis of NMSI incorporates a complex structure, thus accounting for the dependence between the tag and the sentence headed by *qué*, while at the same time incorporating the clausal nature of the tag, which undergoes ellipsis. The main argument to support this hypothesis has to do with interpretation, the focus of the next section.

3.2 The interpretation of NMSI

The phenomenon I analyze here, which, to my knowledge, has gone unnoticed in studies on the subject, is that in NMSI there is a consistent mismatch between form and interpretation. More precisely, sentences in (18), which constitute instances of NMSI, have the same interpretation as the corresponding ones in (17), where the tag contains the (predicate or) whole clause.

- (18) a. *¿Qué te fuiste, a tu casa? =* (17a)
 “What happened/ What did you do, you went home?”
 b. *¿Qué se puso, a llover? =* (17b)
 “What happened/what was it, it started to rain?”
 d. *¿Qué te sientes, incómoda? =* (17d)
 “What’s wrong, you feel uncomfortable?”

In other words, NMSI are interpreted as if a higher constituent, crucially a full clause, appeared in the tag. Importantly, for this interpretation to obtain, the *wh*-element in the main sentence must be neuter *qué*.⁷ All the corresponding matching structures would be unacceptable with the intended interpretation. In (19), also taken from spontaneous Madrid oral speech, the ‘extended’ interpretation is the only possible:

7. Recall that this is the form used when the tag contains predicates or clauses.

- (19) (Spont., Madrid)
- a. *¿Y qué te tienes que quitar toda la ropa?*
 “And what’s wrong/what happens, you have to take off all your clothes?”
[lit. “And what do you have to take off, all your clothes?”]
 - b. *Tienes dos hijos: ¿qué te vas a hacer ahora cargo, de un tercero?*
 “You have two children, what’s next, take charge of a third one?”
[lit. “You have two children, what are you going take charge, of a third one?”]

In sum, my claim is that the tag in NMSI is interpreted as clausal, and that this is because it actually contains a (silent) clause.

One piece of evidence for this claim is that, as has been observed, the tag in split interrogatives may be formed by a disjunction of constituents (as in *what do you want, a car or a bike?*). Interestingly enough, in NMSI the second member of the disjunction whose first member is a DP/PP can be a full sentence, as can be seen in (20), taken from real oral speech:

- (20) a. *¿Qué le trajiste, un regalo o [la llevaste a cenar]?*
lit. “What did you bring her, a present or you took her out for dinner?”
- b. *¿Qué cuesta, un riñón o [se puede pagar]?*
lit. “What does it cost, a kidney or can one pay for it?”
- c. *¿Qué está, lloviendo o [(es que) no tienes ganas de salir]?*
lit. “What is it, raining or is it that you don’t feel like going out?”
- d. *¿Qué lo has resuelto, con calculadora o [(es que) te has vuelto inteligente]?*
lit. “What did you solve it, with a calculator or (is it that) you turned intelligent?”

Given that the members of a coordination or disjunction have to be categorially (and semantically) equivalent, I take the existence of sentences like (20) to show that the element in the tag does not correspond to the focused XP, but it can be a higher (dominating) constituent. My proposal is that the tag in NMSI is the result of an ellipsis process. In addition, as is the case for ellipsis in general (Merchant 2004), a process of reconstruction takes place at LF so that the tag is interpreted as clausal and, consequently, as a wide focus structure. The representation I propose for NMSI is the following:

- (21) $[_{CP} [Qué]_i \dots [_{IP} I \dots [_{FP} t_i [F' \emptyset [_{CP} tag_j [_{IP} \dots [F' I [_{VP} \dots t_j]]]]]]]]]$
-

In (21) there is a node containing both the neuter *qué* and the tag. I propose that this dominating node is related to discourse and contains (as a head) a focus marker. I will label this node as FP/SpeechP and claim that it is the correlate of the discourse projections proposed by monoclausal analysis.

Before going further into the specific proposal, I would like to recall that the interpretation of NMSI is similar to the one obtained in pseudocleft structures. In Fernández-Soriano (2009) it is shown that in the same variety of Spanish there is a type of (non-matching) pseudocleft with the same formal and interpretive properties just described for NMSI. I would like to put forward this idea and develop an analogous analysis.

3.3 Non matching pseudocleft constructions and ‘extended’ interpretation

The similarities between (oral Peninsular) Spanish non-matching pseudoclefts and NMSI are both formal and interpretive. As for the formal part, there is no (Case) matching between the relative pronoun in the (subject) clause and the clefted element (the contrastive focus): instead of the complex relative required in the standard construction, the neuter preform *lo que* (*lit.* ‘the what’) appears. As for the semantic properties, the interpretation of the clefted (focused) constituent after the copula can ‘extend’ to the whole predicate or the whole clause:

- (22) a. *Lo que no puedes prorrumpir es en sollozos.*
 “What you cannot do /what is not acceptable is that you burst into tears”
[lit. ‘The that you cannot burst is into tears’]
- b. *Lo que se puso es a llover.*
 “What happened is that it started to rain” [*lit.* ‘The that it started is to rain’]
 (TVE1/03-09-2006)
- c. Context: After a helicopter accident a journalist asked the president of the Spanish government (M. Rajoy) if he was scared of flying:
Miedo no, lo que no he vuelto a montar es en un avión pequeño.
 “Not really scared, but (what happens/the truth is that) I have not got into a small plane since then” [*lit.* ‘Fear not, the that I have not got again is in a small plane’]
 (M. Rajoy. Telecinco News, 2008)

In Fernández-Soriano (2009) this fact is analyzed as a case of covert contrastive focus extension.⁸ For the focus to propagate, a process of reconstruction (similar to the one proposed for ellipsis contexts) must take place. Based on Kiss’s (1998) approach, the structure of (22a) would be (23).

- (23) [_{CP/IP} [_{CPj} Lo que no puedes prorrumpir [_T t_i es_k [_{FP} [_{PP} en sollozos_i t_k [_{CP} ... t_j]]]]]]]

The hypothesis I put forward is that NMSI have a similar underlying structure. In the next section, I give support for the proposed reconstruction process which explains the obtained interpretation just described.

8. It is usually assumed that it is informative/neuter focus that propagates. See Zubizarreta (1998).

4. Reconstruction: Constraints on movement and idioms

It has been pointed out that NMSI are more restricted than MSI with respect to constraints on movement. Thus, operators such as negation and *wh*-elements seem to intervene between the main sentence and the tag,⁹ as shown by the following sentences (ungrammatical with the intended interpretation).

- (24) a. [?]¿A quién sabes cuándo visitaron, a Pedro?
 “To whom do you know when they visited, Pedro?”
 b. *¿Qué sabes cuándo visitaron, a Pedro?
- (25) a. ¿Adónde no quería ir Juan, a la playa?
 “Where didn’t John want to go, to the beach?”
 b. *¿Qué no quería ir Juan, a la playa?

I contend that this fact is expected since the required reconstruction process after ellipsis is blocked.

Maybe the most important piece of evidence in favor of reconstruction is that NMSI may involve elements that cannot be affected by movement processes, such as idioms. Structures in (26), where the phrases in the tags are stranded form idioms, are perfectly acceptable in the described dialect:

- (26) a. ¿Qué metiste, la pata?
 “What happened, you put the foot on your mouth?” [*lit.* “What did you put in, the paw?”]
 b. ¿Qué cuesta, un riñón?
 “What’s wrong, is it too expensive?” [*lit.* “What does it cost, a kidney?”]
 c. ¿Qué lo has mandado, a la mierda? (Spont., Madrid)
 “What did you do, send him to hell?” [*lit.* “What did you send him, to the shit?”]
 d. *Y yo ¿qué me chupo, el dedo?* (Spont., Madrid)
 ‘What do you think, do you think I am dumb?’ (‘What am I, dumb?’) [*lit.* “What do I suck, my finger?”]¹⁰

9. López Cortina (2009), for example, claims that this is due to *qué* adjunction to CP.

10. The idioms involved in these sentences are the following:

- (i) *Meter la pata* = to put the foot on one’s mouth
Costar un riñón = to be unaffordable
Mandar a alguien a la mierda = to send someone to hell
Chuparse el dedo = to be dumb/childish.

The only possible reading for these sentences is the ‘extended’ one, where the tag is interpreted as (the predicate or) the whole clause containing the idiom. Again, the equivalent structures with matching *wh*-elements are impossible, as expected. My claim is that reconstruction of the idiom at LF would explain the possibility above.¹¹ These data may be seen as instantiating so called Repair Effects, which are directly related to reconstruction (see Merchant (2004) for other Repair cases in Japanese).

My claim is that in NMSI there is an ellipsis process. According to Merchant (2004), ellipsis is licensed by the presence of an E feature in some heads, allowing its complement to be erased in PF. In this case, I suppose that the relevant E feature is in FP/CP. The reconstruction process that takes place at LF is necessary to satisfy e-givenness and F-closure. The former states that only constituents that have a salient antecedent in discourse can undergo ellipsis; the latter requires that the antecedent of the elided constituent contain only thematic material. In the next section, I will show that the tag is a fragment answer to the main *wh*-question. Fragment answers have been claimed to undergo a process of reconstruction (Merchant 2004) and this is, I contend, what happens in NMSI.

5. Ellipsis: NMSI and fragment answers

Merchant (2004: 716) defines fragment answers as “non-sentential utterances with propositional meaning and the force of assertions”. His proposal is that the fragment moves to a clause-peripheral position, FP (identified with Rizzi’s FocusP), and then ellipsis of the clause itself applies. The head F contains the E feature that triggers non-pronunciation of TP. Examples in (27), from Merchant (2004: 673), include the corresponding ‘complete’ structure:¹²

11. Note that if the idiom contains a ditransitive construction, the two objects cannot be separated:

- (i) a. *¿Qué le pides peras, al olmo? (¿Qué le pides, peras al olmo?)
“What do you ask pears to the elm (What is it, you’re asking for something impossible?)”
- b. *¿Qué le diste con la puerta, en las narices? (¿Qué le diste, con la puerta en las narices?)
“What did you hit him with the door, on the nose (What happened, you threw him away?)”
- c. *¿Qué tenía un as, en la manga? (¿Qué tenía, un as en la manga?)
“What did he have, an ace in the sleeve (what was it, he had an ace up his sleeve?)”

All the structures above are acceptable if the complex object as a whole is moved, leaving the verb (which is also part of the idiom) behind.

12. See also Van Craenenbroeck (2010), among others.

- (27) a. *What does Bush want to do to Iraq?*
Attack it. (Bush wants to attack it.)
- b. *What’s left for me to eat?*
Some turkey. (There’s some turkey.) (Merchant 2004: 673)

The fragments in cases like (27) show grammatical dependencies/connectivity effects identical to those of the fragment’s correlate in a non-elliptical sentential structure. Crucially, that the tag in NMSI has all these properties: it is an XP with the force of a full (yes/no) question, where the relevant fragment moves to the left periphery and the non-focal IP is elided. Merchant shows that fragment answers (as ellipsis processes in general) undergo a reconstruction process. My proposal is that, in LF, NMSI’s tags also contain a (reconstructed) full clause.

In fact, like fragment answers with ellipsis, NMSI show connectivity effects for binding and scope. Arregi (2010) describes these phenomena for MSI, where the tag behaves as if it were inside the *wh*-clause with respect to Binding Conditions. The following examples show that NMSI structures share this property:¹³

- (28) a. *¿Qué está hablando Juan, consigo mismo?*
lit. “What is John talking to, himself?”
- b. *¿Y qué llevaba cada una, a su hijo?*
lit. “And what was each one carrying, her child?”
- c. *¿Qué pro_i se queda, en casa de Juan_{i/j}?*
lit. “What is he staying, at Juan’s house?”

In addition, as fragment answers, NMSI are sensitive to islands. According to Merchant (2004), island sensitivity is due to the impossibility of the required reconstruction process to take place. It can be seen that this is exactly what happens with split interrogatives:

13. In NMSI the corresponding structure without ellipsis of non-focal material in the tag is impossible (i). This was noted by Arregi (2010) for matching split interrogatives, and Merchant (2004: 695f) shows that it also happens with fragment answers, and provides examples like (ii) from (American) English:

- (i) a. *¿A CASA, (has dicho que) vas?
 Home you-said that you-are-going
- b. *¿CON UNA CALCULADORA, lo has resuelto?
 With a calculator it you.have solved
- c. *¿A TU JEFE, se lo vas a decir?
 To your boss it you.are going to tell
- (ii) a. – *What has John done?*
 – *Broken the base.*
- b. *BROKEN THE VASE, John has.

According to the author, there are constraints that can be satisfied by non-pronunciation.

- (29) a. *¿Qué te arriesgaste para que salvaran, a tu padre?
 What CL you.risk for that they.save, to your father
 b. *¿Qué te gustaba el candidato que citara, a Chomsky?
 What CL you.like the candidate who quoted, to Chomsky?

NMSI contrast with MSI in that only the former shows sensitivity to negation:

- (30) *¿Qué no se lo diste, a Juan?
 What not CL CL you.give, to Juan?
 (vs. ¿A quién no se lo diste, a Juan?)

The only apparent exception to connectivity effects is (lack of) Case matching between the *wh*-element and the tag. Merchant (2004) deals with similar cases in Japanese for focused DP's in cleft-like structures and shows that these are also subject to ellipsis. In what follows, the proposal for NMSI will be formalized.

6. Ellipsis in (Right) dislocation structures

Let us first consider 'Contrastive Left-Dislocation' (CLD), studied by Ott (2014). CLD is found in most Germanic languages and differs from regular CLLD in that the left dislocated element typically has the pragmatic function of a contrastive topic or a focus. As Ott (2014) shows, what appears in initial position in these structures is a remnant of clausal ellipsis. (31) shows an example and the corresponding structure, where the dislocated phrase has A'-moved to the edge of CP₁, enabling constituent deletion of the remainder of the clause. In the mapping to phonetic form, the material to the right of the fronted XP in CP₁ is deleted. Like NMSP, CLD is a root phenomenon. Ott (2014) does not argue for (or against) the option of a structural connection between the two clauses.¹⁴

- (31) *Den Peter, den habe ich gesehen*
 Peter, him have I seen
 [_{CP1}[den peter]_i [~~habe ich t_i gesehen~~]] [_{CP2} den_k habe ich t_k gesehen]
 (Ott 2014: 270)

In Ott & de Vries (2014) a 'Movement cum deletion' approach is also proposed for Right dislocation. In this case, ellipsis occurs in CP₂:¹⁵

14. Still, he mentions that "there seems to be little evidence for d[islocated]XP and host clause forming a single-rooted syntactic object" (Ott 2014: 281).

15. Truckenbrodt (2012) also argues that right dislocated XPs are remnants of ellipsis.

- (32) *Ich habe ihn gestern noch gesehen, den Peter.*
 I have him yesterday still seen the Peter
 ‘I saw Peter just yesterday.’
 [_{CP1} Ich habe ihn gestern noch gesehen] [_{CP2} [den Peter]_i [~~habe ich gestern noch~~
~~t_i gesehen~~]]

I propose a similar analysis for NMSI, with the difference that I contend that, as has been shown, there is a dependence relation between the main *wh*-clause and the clause containing the dislocated XP.¹⁶ The proposed structure for (33a) would be like (33b).

- (33) a. *¿Qué has visto, a tu padre?*
lit. ‘What have you seen, your father?’
 b. [Qué has visto [FP [a tu padre]_i [~~has visto~~ t_i]]]

As said above, a sentence like (33) is interpreted with the *wh*-element (*qué*) scoping over FP, although it is only the fronted focus that is spelled out after ellipsis. This is a consequence of a reconstruction process of the elided material into FP. I would like to mention that the same reasoning based on ‘covert reconstruction’ into ellipsis sites is used by Den Dikken, Meinunger & Wilder (2000) and by Kluck (2011) to account for connectivity effects in (certain types of) pseudoclefts and sentence amalgams, respectively.

In the next section, I turn to the obligatory choice of the neuter *wh*-element *qué* and relate it to similar interrogative structures with so-called ‘Scope Markers’.

7. The non-matching property: The neuter *qué* and Scope Markers

The goal of this section is to account for the ‘formal mismatch’ shown by NMSI, that is, the presence of the dummy (neuter) form *que*, instead of the complex *wh*-phrase that would correspond to the XP in the tag. First of all, let us recall that *qué* is the operator scoping over predicates and clauses, so what NMSI show is a sentential *wh*-operator in the main clause binding a phrasal (DP/PP) element

16. One anonymous reviewer notes that there is a contrast between matching (ia) and non-matching (ib) split questions, and points out that this may constitute further evidence for the claim that the structure of matching questions can correspond to a sequence of individual clauses, whereas in non-matching ones there is a dependency relation, such as the one represented in (21).

- (i) a. *Pregunta que a dónde fuiste, que si a casa*
He asks where you went, that if home
 b. **Pregunta que qué fuiste, que si a casa*
 He asks that what you went, that if home

in the tag. Crucially, the tag is interpreted as clausal. In order to account for this fact, consider so called ‘Scope Marker Constructions’ (SMC), found in German, Hungarian and Hindi, among other languages (see Horvath 2000 and references therein). SMC can be described as complex interrogative constructions, where a ‘minimal’ *wh*-word (equivalent to *what*) appears in the main clause (the SM), while a complex *wh*-element heads the embedded structure. (34) is an example, with the corresponding interpretation:

- (34) *Was denkt sie [wen_i [Fritz eingeladen hat t_i]]?*
 What thinks she whom_{ACC} Fritz invited has
 “What do you think? who has Fritz invited?”

(Lutz, Müller & von Stechow 2000: 5)

I argue that NMSI are parallel to SMC like (34). In both cases there is a biclausal structure, with movement of a constituent to the left periphery of the embedded clause, while in the root clause an operator appears. This operator does not match the features of the moved constituent but is a dummy *wh*-operator, and has ‘sentential scope’.

In what has been called ‘indirect dependency’ approach to SMC, the *wh*-scope marker in the root clause is an ordinary *wh*-argument that quantifies over propositions rather than over individuals (that is, the quantifier equivalent to *what* binds a propositional variable).¹⁷ As the glosses indicate, SM constructions are interpreted as ‘sequences of questions’. The relevant semantic devices for the interpretation of these constructions are claimed to be independently needed. In particular, the second sentence is interpreted as a restriction of the *wh*-object yielding a question denotation (Dayal 1994). My hypothesis is that this analysis can account for the presence of neuter *qué* in NMSI. As proposed for SMS (Dayal 1994), the wide scope interpretation obtains as a result of the integration of the denotation of a full question (the second clause of the sequence) into the *qué*-clause as the restriction of the existential quantifier, binding a propositional variable. So the scope marker is a correlate of the CP. As for the dependency relation between the two clauses, Dayal (1994) proposes two distinct hierarchical positions of the subordinate CP, depending on the language type: IP-adjoined position vs. complement position. IP adjunction seems to be the best analysis for NMSI.

For what has been seen in the previous description, there are (at least) two differences between the two constructions under study. The first one is that, instead of

17. Alternatively, a semantically inert A-bar expletive element, rather than an instance of the normal (contentful) *wh*-quantifier, has been proposed for these structures in so called ‘direct dependency approach’. See Lutz et al. (2000) for discussion. In this paper the “indirect dependency approach” will be adopted.

a *wh*-question, NMSI involve yes/no questions. For this reason, a (non-matching) dislocated/focalized phrase, as opposed to a *wh*-phrase, appears in Spec FP. Nevertheless, as Horvath (2000) shows, this possibility does exist for SMC in some languages, like Hindi. According to Brandner (2000), this is due to the fact that in Hindi the *wh*-element appearing in the main clause in SMC serves as a typer for yes/no questions (this element actually marks simple yes/no questions in this language). The possible answers for NMSI give values for the yes/no question, and not for the *wh*-element in CP1. Dayal (2000: 171) claims that, “in doing so, they embed the proposition corresponding to CP2 as the complement of the verb in CP1”. In Spanish there are structures like the ones in (35), which are yes/no questions (as shown by the tag in parentheses) formed by the scope marker and the tag (with ellipsis):

- (35) *¿Qué, vas para casa (, no)?*
lit. “What, you are going home (aren’t you)?”
*¿Qué, quieres más (, eh)?*¹⁸
lit. “What, you want some more, (hey)?”
¿Qué, trabajando (, verdad)?
lit. “What, working (, right)?”
¿Qué, con hambre (, no)?
lit. “What, with hunger/hungry (aren’t you)?”

The other difference is that, as we saw, in NMSI a process of ellipsis of non-focal material takes place, followed by reconstruction, accounting for their unexpected properties. Since Scope Marker Constructions do not involve ellipsis (or reconstruction), one does not expect ‘extended’ interpretations.

8. Conclusions

The goal of this paper was to account for ‘Non-matching’ Split Interrogatives of (spoken Peninsular) Spanish, where the dummy (neuter) *qué* ‘what’, instead of a matching *wh*-word, heads the matrix clause. In so doing a property unnoticed in

18. In Catalan it is well-known that the particle *que* (different form the *wh*-pronoun) marks yes/no questions:

- (i) *Que en vols més?*
 That CL you.want more
 “Would you like some more? (is it that you want some more?)”
Que tanqueu ara?
 That you.close now
 “Are you closing?”

previous studies has been analyzed: in NMIS there is also a form-meaning mismatch whence the tag, despite the fact that it is formed by a DP/PP, is interpreted as clausal. NMSI also have the unexpected property of allowing idioms to be affected by a syntactic (movement) operation. I proposed a biclausal analysis involving an ellipsis process similar to the one taking place in fragments (Merchant 2004), followed by reconstruction. This accounts for the apparent ‘extension’ in the interpretation of the tag. As for the syntactic analysis, I propose NMSI to instantiate cases of Scope Marker Constructions involving yes/no questions thus accountings for the obligatory presence of the neuter *wh*-form.

Acknowledgements

This work was initiated during my stay at the State University of New York at Stony Brook. I deeply thank Paco Ordóñez, Cristina Tortora and Richard Larson for their support and their useful comments and suggestions. I also thank Francesc Roca for interesting discussion. I thank Violeta Demonte for generous comments on earlier versions of this work. Errors and mistakes are all my own.

Funding

The research behind this work has been supported by the project FFI2017-87140-C4-4-P, from the Spanish Ministerio de Industria, Economía y Competitividad.

References

- Arregi, Karlos. 2010. Ellipsis in split questions. *Natural Language & Linguistic Theory* 28(3). 539–592. <https://doi.org/10.1007/s11049-010-9097-x>
- Brandner, Ellen. 2000. Scope Marking and Clausal Typing. In Uli Lutz, Gereon Müller & Arnim von Stechow (eds.), *Wh-scope marking*, 45–77. Amsterdam: John Benjamin. <https://doi.org/10.1075/la.37.03bra>
- Camacho, José. 2002. Wh-doubling: Implications for the syntax of wh-movement. *Linguistic Inquiry* 33. 157–164. <https://doi.org/10.1162/002438902317382206>
- Contreras, Joan M. & Francesc Roca. 2007. D’oracions interrogatives: les interrogatives escindides. *Caplletra* 42. 145–184.
- Van Craenenbroeck, Jeroen. 2010. *The Syntax of ellipsis. Evidence from Dutch dialects*. Oxford: Oxford University Press.
- Dayal, Veneeta. 1994. Scope Marking as indirect wh dependency. *Natural Language Semantics* 2. 137–170. <https://doi.org/10.1007/BF01250401>

- Dayal, Veneeta. 2000. Scope Marking: Cross-linguistic variation in indirect dependency. In Uli Lutz, Gereon Müller & Arnim von Stechow (eds.), *Wh-scope marking*, 157–194. Amsterdam: John Benjamin. <https://doi.org/10.1075/la.37.07day>
- den Dikken, Marcel. 2006. *Relators and linkers. The syntax of predication, predicate inversion, and copulas*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/5873.001.0001>
- den Dikken, Marcel, André Meinunger & Chris Wilder. 2000. Pseudo-clefts and ellipsis. *Studia Linguistica* 54(1). 41–89. <https://doi.org/10.1111/1467-9582.00050>
- Fernández-Soriano, Olga. 2009. On the nature of covert operations. The case of focus in Spanish pseudoclefts. In Enoch Oladé Aboh, Elisabeth van der Linden, Josep Quer & Petra Sleeman (eds.), *Romance languages and linguistic theory*, 87–109. Amsterdam: John Benjamins. <https://doi.org/10.1075/rllt.1.05sor>
- Grimshaw, Jane. 2005. *Words and structure*. Stanford, CA: CLSI Publications.
- Horvath, Julia. 2000. On the syntax of “wh-scope-marker” constructions: Some comparative evidence. In Uli Lutz, Gereon Müller & Arnim von Stechow (eds.), *Wh-scope marking*, 271–316. Amsterdam: John Benjamin. <https://doi.org/10.1075/la.37.11hor>
- Irurtzun, Aritz. 2017. On the nature and distribution of split wh-questions in Basque. Paper presented at the 27th Colloquium on Generative Grammar, Madrid, Spain, May 17–19.
- Kiss, Katalin É. 1998. Identificational focus versus informational focus. *Language* 74. 245–273. <https://doi.org/10.1353/lan.1998.0211>
- Kluck, M. Elisabeth. 2011. *Sentence amalgamation*. Groningen: LOT, Netherlands Graduate School of Linguistics.
- Lorenzo, Guillermo. 1994–95. *Qué* expletivo en preguntas dislocadas. *Archivum Ovetensis* XLIV–XLV. 423–446.
- López Cortina, Jorge. 2009. Split questions, extended projections, and dialect variation. In Joseph Collentine, Maryellen García, Barbara Lafford, & Francisco Marcos Marín (eds.), *Selected Proceedings of the 11th Hispanic Linguistics Symposium*, 219–230. Somerville, MA: Cascadilla.
- Lutz, Uli, Gereon Müller & Arnim von Stechow. 2000. *Wh-scope marking*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.37>
- Merchant, Jason. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27. 661–738. <https://doi.org/10.1007/s10988-005-7378-3>
- Munaro, Nicola & Jean-Yves Pollock. 2005. ‘Qu’est-ce-que.qu)-est-ce-que?’ A case study in comparative Romance interrogative syntax. In Guglielmo Cinque & Richard S. Kayne (eds.), *The Oxford handbook of comparative syntax*, 542–606. Oxford: Oxford University Press.
- Ott, Denis. 2014. An ellipsis approach to contrastive left-dislocation. *Linguistic Inquiry* 45(2). 269–303. https://doi.org/10.1162/LING_a_00155
- Ott, Denis & Mark de Vries. 2014. Right-dislocation as deletion. *Natural Language & Linguistic Theory* 34(2). 641–690. <https://doi.org/10.1007/s11049-015-9307-7>
- Truckenbrodt, Hubert. 2012. On the prosody and syntax of right-dislocation. Paper presented at Deutsche Gesellschaft für Sprachwissenschaft (DGfS), Frankfurt, Germany, March 7–9.
- Uriagereka, Juan. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26. 79–123.
- Zubizarreta, M. Luisa. 1998. *Prosody, focus and word order*. Cambridge, MA: MIT Press.

Differential object marking and scale reversals

Monica Alexandrina Irimia and Anna Pineda

University of Modena and Reggio Emilia / Sorbonne Université

This paper focuses on some problematic aspects of the diachrony of differential object marking in Old Catalan and Old Romanian (11th to 17th centuries). Corpus data from both languages reveal two unexpected facts: (i) there is a prominence of 3rd person to the exclusion of 1st and 2nd person, contrary to what the Animacy/Person scale would predict; (ii) differential marking appears to be present on nominals (especially proper names), to the exclusion of pronouns, this time contrary to the Specificity/Definiteness Scale. The account we propose for these types of scale reversals builds on the idea that languages can have more than one differential object marking strategy, as well as more than one type of structure for pronouns and animate nominals. Moreover, the co-existence of various mechanisms for nominal licensing can explain why, in some instances, classes lower down the hierarchies can get signaled to the exclusion of higher ones.

Keywords: differential object marking, animacy, scale, reversals, Old Catalan, Old Romanian

1. Introduction

Recent studies have strengthened the important role animacy – and discourse related specifications, more generally – play in the inner workings of the human language computational system. A well-known exemplification comes from differential object marking (DOM), a robust morphosyntactic means of signaling out distinct classes of direct objects (Aissen 2003, a.o.). In many Romance varieties, a generalization is that the ‘higher animates’ are the most likely to receive differential marking. What counts as the most likely ‘higher animates’ is dictated by rankings of the type put forward in the typological literature under the rubric of scales (Silverstein 1976; Aissen 2003; Næss 2004; Dalrymple & Nikolaeva 2011, a.o.). In this paper, we focus on the ‘Animacy/person scale’ and the ‘Specificity/definiteness scale’, illustrated below.

- (1) Types of scales
 - a. *Animacy/person*: 1/2 > 3 > proper name > human > animate > inanimate
 - b. *Specificity/definiteness*: pronoun > name > definite > specific indefinite > non-specific

Some accounts assimilate scales to descriptive statements with regards to linguistic properties, which appear to be typologically robust, although not exceptionless (see the discussion in Carnie 2005; Kiparsky 2008; Avram and Zafiu 2017, a.o.). Under other formalizations (for example Kiparsky (2008)), scales are not a uniform class; while some scale hierarchies encode true universals and have the power to constrain synchronic grammars and language change, others simply indicate typological generalizations and are just the result of regular and typical paths of change. Kiparsky (2008) classifies the scales we are addressing here as true universals. Yet other hypotheses explicitly hold that scales carry non-trivial explanatory power in the sense that they derive implicational universals, and regulate patterns of language change (von Heusinger, Klein & de Swart 2008, a.o.). Along this line, Romance differential object marking is claimed to have systematically arisen with 1st and 2nd person pronouns, which carry the ‘highest specification of animacy’, and then slowly extended to other classes of nominals (Rohlf 1971, 1973; Roegiest 1979; Sornicola 1997, 1998; Bossong 1991; Leonetti 2003, 2008; Mardale 2015, a.o.). Against this background, one intriguing question is whether there are instances of language change or synchronic processes where hierarchical generalizations are not respected and, if so, how they should be explained.

It is precisely this aspect we examine here. We focus on the diachrony of differential object marking in Old Catalan and Old Romanian (11th to 17th centuries); we illustrate two contexts where canonical scale rankings appear not to be respected. More specifically, corpus data from Old Catalan (OC) and Old Romanian (OR) permit two unexpected conclusions. On the one hand, following the observations in Irimia & Pineda (2019), there is a prominence of 3rd person, to the exclusion of 1st and 2nd person, contrary to what the Animacy/Person scale in (1a) predicts. On the other hand, differential marking appears to be present on nominals (especially proper names), to the exclusion of pronouns, this time contrary to the Specificity/Definiteness Scale in (1b). We propose an analysis for these contexts which we label as ‘scale reversals’. The gist of our account is that languages can have more than one differential object marking strategy, as well as more than one type of structure for pronouns and animate nominals. The co-existence of various structural mechanisms for nominal licensing can explain why in some instances, classes lower down the hierarchies can get signaled to the exclusion of the higher ones. The larger picture conclusion that emerges is that ‘truly universal’ scales can be ‘reversed’ but not necessarily because the rankings they encode might not be universal, but because languages can contain more than one grammatical strategy to license nominal categories (among other explanations).

The structure of the paper is as follows. In § 2, building on the findings in Irimia & Pineda (2019) we discuss data from OC and OR. These signal the prominence of 3rd person via a reversal of the Animacy/Person scale. We also introduce the other scale reversal, namely the presence of DOM with nominals to the exclusion of pronouns. Section 3 proposes an analysis that unifies the two reversal contexts, whose main tenet is that there can be more than one (differential object) marking strategy active in a language at the same time. Section 4 presents the conclusions.

2. DOM in Old Romanian and Old Catalan

As mentioned above, the established wisdom with respect to Romance DOM is that it arose with 1st/2nd person pronouns and then extended to full nominals. Old Spanish appears to respect this observation (see especially von Heusinger & Kaiser 2005: 35–36, 41, a.o.). On the basis of this remark, it is very often assumed that presence of DOM on 3rd person to the exclusion of 1st/2nd person is, in fact, impossible not only diachronically but also synchronically.

In OC, however, it does not seem to be the case that DOM consolidated first with 1st/2nd person and then extended to 3rd person. Irimia & Pineda (2019) examined various texts from the 11th to the 16th centuries and provide patterns like the ones in (2). What is puzzling here is that the 3rd person pronoun shows DOM, whereas the 1st and 2nd person ones unexpectedly lack it.

(2) (Curial e Güelfa, 15th c.)

- a. *vós havets honrat a ell.*
you.2PL¹ have.2PL honoured DOM he
“You have honored him.”
- b. *aquella senyora, qui mira nosaltres ...*
that lady who look.3SG we
“That lady, who watches us ...”
- c. *¿què ha mogut tu e ton companyó a ...?*
what has moved you and your companion to
“What compelled you and your mate to ...” (Irimia & Pineda 2019: 5)

In fact, in some texts the proportion of 1st/2nd person pronouns signaled by DOM goes down even to 0%, while that of differential marking on 3rd person pronouns reaches 100%. We summarize some statistical generalizations in Table 1 below. Note that while the number of tokens can indeed be low, due to the nature of the texts, what matters is the overwhelming general tendency of DOM with 3rd person.

1. We have followed the Leipzig glossing conventions for abbreviations.

Table 1. DOM on pronouns in Old Catalan (from Irimia & Pineda 2019: 2–3, Tables 1–5)

Old Catalan text	1st/2nd person	3rd person
<i>Clams e crims</i> (13th c.) [30.000 words]	0/1 (0%)	3/4 (75%)
<i>Novel-letes Hongria</i> (14th c.) [9.000 words]	0/2 (0%)	3/9 (33.3%)
<i>Clams e crims</i> (14th c.) [30.000 words]	0/3 (0%)	2/2 (100%)
<i>Epistolari València</i> (14th c.) [30.000 words]	5/14 (35.8%)	4/5 (80%)
<i>Curial e Güelfa</i> (15th c.) [30.000 words]	5/9 (55.5%)	5/6 (83.3%)
<i>Epistolari València</i> (15th c.) [30.000 words]	3/12 (25%)	2/2 (100%)

A similar situation occurs in Old Romanian, as has initially been observed by von Heusinger & Onea Gáspár (2008). In an examination of some of the first Bible translations into Romanian,² they observed that 3rd person pronouns were differentially marked to an overwhelmingly higher degree and in a more stable way than 1st and 2nd person pronouns, as seen in Table 4 below. But this is exactly the opposite of what the scales predict. Given the hierarchical implications, DOM is not expected to skip 1st and 2nd persons if 3rd person is signalled morphologically. Below we illustrate examples from a 16th century text. We notice here the DOM preposition *pe* (“on”) under its variants. There is a contrast between a 3rd person pronoun, which shows both DOM as well as (accusative) clitic doubling as in (3a), and a 1st person pronoun seen without DOM, as in (3b):

(3) Tetraevanghel – Coresi (CC²)³

- a. *Surpa-l-va* *pre elu.*
 destroy-CL.ACC.M.3SG-FUT.3SG DOM he
 “He will destroy him.” (CC².1581: 23)

2. As discussed by various authors, the use of DOM in Old Romanian appears to show differences when it comes to translations as opposed to original texts (Stan 2013, a.o.). There are also differences between DOM conditions of use in the first part of the 16th century and texts of the end of the 16th century (Pușcariu 1921/1922, a.o.). Other researchers have presented a more refined picture of the behavior of 3rd person pronouns (Avram & Zafiu 2017; Hill & Mardale 2021, a.o.), revealing important degrees of variation across texts. What matters, however, is that there seem to be (early) corpora in which 3rd person is stronger than 1st and 2nd persons as a DOM trigger.

3. Full corpus names are listed at the end.

- b. *Va vindeca noi.*
 FUT.3SG heal we
 “He will heal us.” (CC².1581: 20)

von Heusinger & Onea Gáspár’s (2008) results, supplemented with the analysis of other texts in Irimia & Pineda (2019), clearly mirror the Old Catalan picture, raising significant questions about the nature of DOM and whether it is indeed constrained by the scales.

2.1 Pronouns and DPs

Besides the contexts of 3rd vs. 1st/2nd person pronouns, which are analysed in detail in Irimia & Pineda (2019), we also focus on another counterexample to scales in OC and OR: in the two languages, (1st and 2nd person) pronouns can appear without DOM, while other DPs show DOM. The earliest Catalan texts also demonstrate that *a*-marking did not necessarily consolidate first with strong pronouns and then extended to DPs. In particular, for Catalan we refer to DPs formed from a proper name preceded by the ‘treatment’ form (*e*)*n/na* (which would evolve into a personal definite article) or preceded by the sequence *el dit* ‘the above-mentioned’. See the contrast in (4):

- (4) a. *darem a aquels [l]icència de peynorar vós*
 give.FUT.1PL to them permit to fine.INF you.2PL.HONORIFIC(=SG)
 “We will give them permit to fine you.”
- b. *com en Ca[ste]let, saig, volgués peynorar a· n*
 since the Castelet, executioner, want.SBJV.PST.3SG fine.INF DOM the
Ramon Sanç, lo dit Ramon dix a aquel que
 Ramon Sanç the mentioned Ramon tell.PST.3SG to that.one that
no·l peynoràs,
 no him.ACC fine.SBJV.PST.3SG
 “Since Castelet, executioner, wanted to fine Ramon Sanç, the above-mentioned Ramon told to that one not to fine him.”
 (Clams e crims, 13th c., page 33 line 19 and page 54 line 30)

More specifically, in the text the examples in (4) are taken from, the percentages of DOM with the two types of objects are as in Table 2. What we conclude from these statistics is that DOM was not always salient on 1st and 2nd person pronouns, once again contrary to the predictions of the scales.

Table 2. DOM on pronouns and DPs with proper names in one of the OC texts

Proportion of DOM in a text from Old Catalan	1st/2nd	3rd	DPs containing proper names
<i>Clams e crims</i> (13th c.) [30.000 words]	0/1 (0%)	3/4 (75%)	16/77 (20.8%)

Turning to Old Romanian, we observe the same picture with even more robust evidence. In an independent study, Nicula Paraschiv (2016) also notes that at least for the 1st person pronoun, its DOM-less instances outrank the DOM ones in some 16th century texts.

Table 3. *P(re)* marking of the personal DO – 16th c. (Nicula Paraschiv 2016)

Text	Form	+ <i>pre</i>	- <i>pre</i>
PO. 1582	mene/menre (I.ACC)	12	16

Moreover, with respect to the presence of DOM on DPs to the exclusion of 1st/2nd person (and 3rd person) pronouns, the same results were in fact obtained by von Heusinger & Onea Gáspár (2008). Their findings are summarized in Table 4 below. Note that the presence of DOM is exceptionless on proper names (for animates) in these texts, in fact over-ranking even 3rd person pronouns.

Table 4. Direct object marking with *p(re)* in the 16th century Romanian (von Heusinger & Onea Gáspár 2008: 77, Table 1)

	Me/you	Other pronouns (one instance of reflexive)	Prop. name	Def. NP	Indef.NP	Total
+ <i>pe</i>	5	33	3	9	1	51
- <i>pe</i>	5	1	0	36	9	51
Total	10	34	3	45	10	102
%+ <i>pe</i>	50%	97%	100%	20%	10%	50%

Similar observations are to be found in Nicula Paraschiv (2016: 132), who has noticed that proper names can appear without DOM only very rarely. Avram & Zafiu (2017: 36), in fact, mention the plausibility of “an earlier stage [of Romanian, our note] when proper names might have been more robustly *pe*-marked than definite pronouns”, according to the scale in (5):

- (5) Proper names > definite pronouns > definite DPs > indefinite DPs
(Avram & Zafiu 2017: i)

Also note that a third study into OR DOM has obtained the same results. Hill & Mardale (2019, 2021) have examined a larger variety of OR texts and their findings are summarized in Tables 5 and 6. Here we see, again, a higher percentage with DOM on nominals when compared with pronouns, at least in some of the texts.

Table 5. Direct object marking with *p(r)e* in the 16th century Romanian (Hill & Mardale 2021: Table 3)

Operation	CEv		PO		Dî	
	Pron.	Nouns	Pron.	Nouns	Pron.	Nouns
DOM-p	271; 48.91%	90; 42.05%	214; 49.76%	245; 62.02%	17; 64%	56; 76.71%

Table 6. Direct object marking with *p(r)e* in the 17th and 18th centuries Romanian (Hill & Mardale 2021: Table 4)

Operation	17th century				18th century			
	Varlaam		Ureche		Ivireanu		Neculce	
	Pron.	Nouns	Pron.	Nouns	Pron.	Nouns	Pron.	Nouns
DOM-p	52; 25.12%	240; 94.11%	22; 43.13%	342; 87.91%	87; 43.5%	368; 96.08%	43; 48.31%	526; 90.37%

But this once again raises questions with respect to scales. As we have seen, under the scales all pronouns are predicted to be higher than any other DP. Also note that, as we also mentioned for the first scale reversal, even if occurrences might be low (due to the nature of the texts examined), what matters is the tendency; and that is quite clear not only in OR but also in OC.

In conclusion what we see from both OC and OR data are two conflicting properties: (i) clear sensitivity of DOM to animacy (there are basically no examples with DOM with inanimates in OC texts examined here);⁴ (ii) scale reversals when it comes to hierarchies inside the animate class. In the next section we explore a possible answer to these unexpected patterns.

4. In OR the facts are more complex, but we leave inanimates aside as they are not relevant in OC.

3. Analysis

The empirical landscape we have addressed above has received less attention in the descriptive and theoretical literature, where DOM is taken to respect the Animacy/Referentiality scales. The few cases that have been pointed out to the opposite have been classified as exceptions or unclear examples, if acknowledged at all. Otherwise, von Heusinger & Onea Gáspár are the only account we know of that discussed in details the prominence of 3rd person DOM when compared with that of 1st and 2nd persons. They also put forward a tentative explanation, according to which DOM was needed in OR as a morphological means to signal accusative case. The observation von Heusinger & Onea Gáspár start from is the following. As seen in Table 7 below, OR 1st and 2nd person pronouns have dedicated accusative morphology. This fact is taken to indicate that DOM was only necessary with the 3rd person, due to the nominative – accusative homomorphism in this class:

Table 7. Morphology of personal pronouns in OR (16th century) and OC

	Romanian				Catalan			
	Singular		Plural		Singular		Plural	
	NOM	ACC	NOM	ACC	NOM	ACC	NOM	ACC
1st	eu	mene/menre	noi	noi	jo	mi	nós/nosaltres	nós/nosaltres
2nd	tu	tine	voi	voi	tu	tu	vós/vosaltres	vós/vosaltres
3rd	M elu	elu	ei	ei	ell	ell	ells	ells
	F ea	ea	ele	ele	ella	ella	elles	elles

We follow Irimia & Pineda (2019) in pointing out that this hypothesis suffers from several shortcomings. On the one hand, NOM-ACC disambiguation does not explain why we do not get DOM with 1st and 2nd person objects ‘in the plural’. As we can see in Table 7, there is no dedicated accusative morphology ‘in the plural’, for any person. Yet, DOM is only obligatory with 3rd person, as we have shown above. On the other hand, the data raise yet another problem. OR (like Modern Romanian) also permitted accusative clitic doubling with differential objects. Thus, the corpora provide us with examples like the one in (6), where we can see a 3rd person direct object carrying both DOM as well as clitic doubling. Importantly, the clitic shows up with clear accusative morphology.⁵ If DOM were simply an accusative case disambiguation strategy, it is not clear why it would be needed with nominals that already have accusative case indicated by other means, such as accusative clitic doubling.

5. OR clitics have distinct dative morphology (see the extensive discussion in Pană Dindelegan 2016) and there are no clitics for the nominative.

- (6) *Surpa-l-va* *pre elu.*
 destroy-CL.ACC.M.3SG-FUT.3SG DOM he
 “He will destroy him.” (CC².1581: 23)

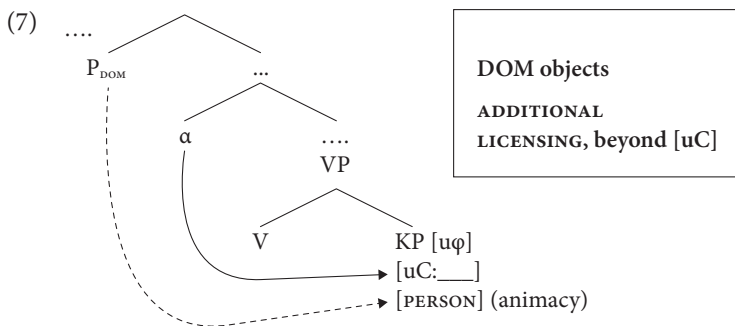
Lastly, if we turn to OC, we are faced with similar problems; tonic pronouns only show distinct accusative morphology in the 1st person singular. The disambiguation hypothesis would predict the obligatoriness of DOM on both 2nd and 3rd person pronouns, contrary to what the examples we have presented indicate. Therefore, the safe conclusion we draw from the data is that the problem is 3rd person⁶ *per se* (as well as other nominals which are highly animate, such as proper names), and not the need to disambiguate an accusative form. We propose that these are not necessarily counterexamples to scales, once we connect DOM to a licensing mechanism which is independent of abstract Case.

3.1 Animacy and person

The most important observation we need to start with is that Romance DOM is broadly regulated by ‘animacy’. Any analysis needs to capture this fact. In the theoretical literature, at least two main directions have been pursued: (a) DOM as a morphological mechanism (i.e., animates are not syntactically different from other objects, they are signaled just in the morphology, see López 2012 or Bárány 2018 for an overview, a.o.); (b) DOM as a syntactic mechanism. Our data are hard to accommodate under a morphological explanation for various non-trivial reasons, such as special types of co-occurrence restrictions reminiscent of the PCC (see Ormazabal & Romero 2013, a.o.), or different syntactic positions for DOM (López 2012, a.o.). All these are basically impossible to derive if differential objects are not assumed to have a distinct ‘syntactic’ structure when compared to other objects.

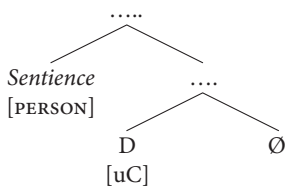
Turning to syntax, in most recent accounts (see López 2012; Ormazabal & Romero 2013, a.o.), DOM is connected to (the spell-out of) a Case licensing mechanism on certain types of objects. We build on this line of research. However, we propose that DOM is linked to a licensing mechanism beyond Case. We derive sensitivity to animacy following recent accounts which connect *grammaticalized* animacy to the presence of a [PERSON] (Cornilescu 2000; Rodríguez-Mondoñedo 2007, a.o.) feature on certain types of nominals, as schematically indicated in (7):

6. Similar scale reversals with prominence of DOM on 3rd person have, in fact, been observed elsewhere. See, for example, the discussion in Dalrymple & Nikolaeva (2011) for data from Finno-Ugric and Indo-Aryan languages.

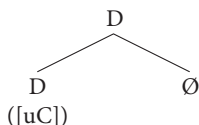


The class relevant to us is 3rd person. As opposed to 1st and 2nd persons, 3rd person pronouns can signal both inanimate and animate entities. The *grammaticalized animate entities* are taken to contain an additional [PERSON] feature, as schematically shown in (8a). This additional specification serves to indicate their discourse saliency, as animates. Thus, 3rd person ‘animate’ pronouns contain an additional [PERSON] feature which requires additional licensing beyond [uC]. Proper names can also contain this additional feature (see Richards 2008, a.o.). We hypothesize that this is due to the need to split animate proper names from inanimate proper names. 1st and 2nd person (and also at least some 3rd person inanimate) pronouns initially contain just a Case feature (as in (9a) and (10a)), connected to the Speaker/Hearer specifications (or other features for 3rd inanimates), whose presence might be signaled by special inflectional morphology (dedicated accusative forms for 1st/2nd pronouns, etc.). The system we obtain contains the parametric options listed below:

- (8) a. Third person animate pronoun

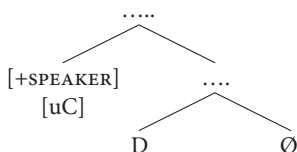


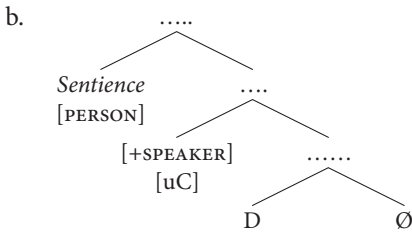
- b. Third person inanimate



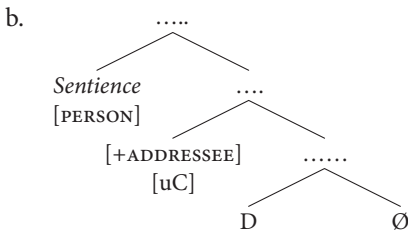
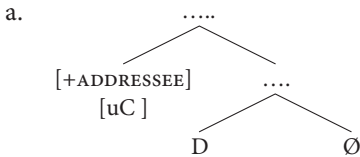
- (9) First person pronoun

- a.





(10) Second person pronoun



In the diachrony of OC and OR, the extension of DOM to 1st and 2nd person can be explained either as a strengthening of animacy (as these pronouns can *only* be animate, and thus a [PERSON] feature has been added to them, as represented in (9b) or (10b)) or as a transition from a nominal licensing strategy based on (morphological) C/case (as in Latin) to a nominal licensing system based on pragmatic specifications, after the [PERSON] feature extended to all animates.

4. Concluding remarks

The data we have examined permit important conclusions about the connection between differential object marking and scales. First, there is the indication that more than one scale might be active at a given moment in a certain system. Second, each of the specifications on a scale can introduce its own hierarchical implications. Third, more than one (nominal) licensing strategy can be active in a language. All these parameters do not necessarily argue against the existence of scales. They support instead a flexible grammatical system where interactions between various types of licensing give the appearance of scale reversals.

The data also provide us with more hints into the divide between true universals and typological generalizations, the latter without the status of true universals. In

a paper addressing the differences between the two classes, Kiparsky (2008) explores a hypothetical counterexample to the universally expected process of coda devoicing; the author notices that across Romance the interaction of independent phonetic and phonological processes applying in a well-determined sequence might, in fact, give rise to ‘coda voicing’. Kiparsky (2008) subsequently concludes that coda devoicing, even if subject to featural hierarchy, cannot be a true universal; scales like the Animacy one must, on the contrary, be universal. What we have illustrated here is that even putatively universal scales can be contradicted; however, these violations might be only apparent. They can be derived from the same principles as the violations of mere typological generalizations – more than one structurally-sensitive process can affect certain well-defined classes in a language at a given moment such that the result might appear to be a contradiction of the application of ‘just one process’.

Acknowledgements

We would like to thank Giuseppe Longobardi, András Bárány, Cristina Guardiano, M. Teresa Espinal, Adam Ledgeway, Rita Manzini, the audiences at DiGS 20 (York University), Going Romance 32 (Utrecht), the LSA 2019 Annual Meeting, as well as several anonymous reviewers for their feedback and discussion on various points addressed in this paper. All errors and mistakes are our own.

Funding

Funding – First author: the ERC grant # 295733 LanGeLin (PI Giuseppe Longobardi) and a research grant offered by Università di Modena e Reggio Emilia. Second author: the postdoctoral research fellowship Juan de la Cierva-*incorporación* and the research project FFI2014-56968-C4-1-P (IJC1-2016-30474, Ministerio de Economía y Competitividad, Spanish Government).

References

- Aissen, Judith L. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21(3). 435–483. <https://doi.org/10.1023/A:1024109008573>
- Avram, Larisa & Rodica Zafiu. 2017. Semantic hierarchies in the evolution of differential object marking in Romanian. In Adina Dragomirescu, Alexandru Nicolae, Camelia Stan & Rodica Zafiu (eds.), *Sintaxa ca mod de a fi. Omagiu Gabrielei Pană-Dindelegan, la aniversare*, 29–42. Bucharest: Editura Universității din București.
- Bárány, András. 2018. DOM and dative case. *Glossa: A Journal of General Linguistics* 3(1). 97. <https://doi.org/10.5334/gjgl.639>

- Bosson, Georg. 1991. Differential object marking in Romance and beyond. In Dieter Wanner & Douglas A. Kibbee (eds.), *New Analyses in Romance Linguistics: Selected Papers from the XVIII Linguistic Symposium on Romance Languages*, 143–170. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.69.14bos>
- Carnie, Andrew. 2005. Some remarks on markedness hierarchies. A reply to Aissen (2003). *Coyote Working Papers in Linguistics* 14. 37–50.
- Cornilescu, Alexandra. 2000. Notes on the interpretation of the prepositional accusative in Romanian. *Bucharest Working Papers in Linguistics* II (1). 91–106., file:///E:/DOM/Romanian/236627088-Cornilescu-00-Notes-on-Th.pdf
- Dalrymple, Mary & Irina Nikolaeva. 2011. *Objects and information structure*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511993473>
- von Heusinger, Klaus & Georg Kaiser. 2005. The evolution of differential object marking in Spanish. In Klaus von Heusinger, Georg Kaiser & Elisabeth Stark (eds.), *Specificity and the evolution /emergence of nominal determination systems in Romance*, 33–69. Konstanz: Universität Konstanz, Fachbereich Sprachwissenschaft.
- von Heusinger, Klaus, Udo Klein & Peter de Swart. 2008. Variation in differential object marking. Paper presented at Workshop on Case variation, Stuttgart, Germany, June 19–20.
- von Heusinger, Klaus & Edgar Onea Gáspár. 2008. Triggering and blocking effects in the development of DOM in Romanian. *Probus* 20. 67–110. <https://doi.org/10.1515/PROBUS.2008.003>
- Hill, Virginia & Alexandru Mardale. 2019. Patterns of differential object marking in the history of Romanian. *Journal of Historical Linguistics* 3. 5.
- Hill, Virginia & Alexandru Mardale. 2021. *The diachrony of differential object marking in Romanian*. Oxford: Oxford University Press.
- Irimia, Monica & Anna Pineda. 2019. Differential object marking and Scales: Insights from Romance diachrony. In *Proceedings of the Linguistic Society of America* 4. 57. 1–15. <https://doi.org/10.3765/plsa.v4i1.4561>
- Kiparsky, Paul. 2008. Universals constrain change: Change results in typological generalizations. In Jeff Good (ed.), *Language universals and language change*, 23–53. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199298495.003.0002>
- Leonetti, Manuel. 2003. Specificity and object marking: The case of *a*. In Klaus von Heusinger & Georg A. Kaiser (eds.), *Proceedings of the Workshop Semantic and Syntactic Aspects of Specificity in Romance Languages*. Arbeitspapier 113, 67–101. Konstanz: Universität Konstanz.
- Leonetti, Manuel. 2008. Specificity in clitic doubling and in differential object marking. *Probus* 20. 33–66. <https://doi.org/10.1515/PROBUS.2008.002>
- López, Luis. 2012. *Indefinite objects: Scrambling, choice functions and differential marking*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9165.001.0001>
- Mardale, Alexandru. 2015. Differential object marking in the first Romanian texts. In Virginia Hill (ed.), *Formal approaches to DPs in Old Romanian*, 200–246. Leiden: Brill. https://doi.org/10.1163/9789004292550_007
- Næss, Ashild. 2004. What markedness marks: The markedness problem with direct objects. *Lingua* 114. 1186–1212. <https://doi.org/10.1016/j.lingua.2003.07.005>
- Nicula Parashiv, Irina. 2016. The direct object. In Gabriela Pană-Dindelegan (ed.), *The syntax of Old Romanian*, 123–143. Oxford: Oxford University Press.
- Ormazabal, Javier & Juan Romero. 2013. Differential object marking, case and agreement. *Borealis: An International Journal of Hispanic Linguistics* 2(2). 221–239. <https://doi.org/10.7557/1.2.2.2808>

- Pană-Dindelegan, Gabriela. 2016. Preface. In Gabriela Pană-Dindelegan (ed.), *The syntax of Old Romanian*, xxiii–xxvi. Cambridge: Cambridge University Press.
<https://doi.org/10.1093/acprof:oso/9780198712350.001.0001>
- Pușcariu, Sextil. 1921/1922. Despre *pre* la acuzativ. *Dacoromania* 2. 565–581.
- Richards, Marc. 2008. Defective agree: Case Alternations, and the prominence of Person. In Marc Richards & Andrej L. Malchukov (eds.), *Linguistische Arbeitsberichte (Volume on Scales)*, Volume 86, 137–161. Leipzig: Universität Leipzig.
- Rodríguez-Mondoñedo, Miguel. 2007. The syntax of objects. Agree and differential object marking. Ph.D. dissertation. University of Connecticut Storrs.
- Roegiest, Eugene. 1979. À propos de l'accusatif prépositionnel dans quelques langues romanes. *Vox romanica* 38. 312–334.
- Rohlf, Gerard. 1971. Autour de l'accusatif prépositionnel dans les langues romanes. *Revue de Linguistique Romaine* 35. 312–327.
- Rohlf, Gerard. 1973. Panorama de l'accusatif prépositionnel en Italie. *Studii și Cercetari Linguvistice* 24. 617–621.
- Silverstein, Michael. 1976. Hierarchy of features and ergativity. In Robert M. W. Dixon (ed.), *Grammatical categories in Australian languages*, 112–71. Canberra: Australian Institute of Aboriginal Studies.
- Sornicola, Rosanna. 1997. L'oggetto preposizionale in siciliano antico e in napoletano antico. *Italienische Studien* 18. 66–80.
- Sornicola, Rosanna. 1998. Processi di convergenza nella formazione di un tipo sintattico: la genesi ibrida dell'oggetto preposizionale. *ACILPR XXII* 2, 419–427.
- Stan, Camelia. 2013. *O sintaxă diacronică a limbii române vechi*. Bucharest: Editura Universității din București.

Sources

OLD CATALAN

Corpus Informatizat del Català Antic (CICA), J. Torruella (ed.), junt amb Manel Pérez Saldanya i Josep Martines. www.cica.cat

OLD ROMANIAN

- CC².1567 Coresi, *Evanghelie cu învățătură*. Ed.: Diaconul Coresi, *Carte cu învățătură* (1581), Volume 1, *Textul*, ed. S. Pușcariu, Al. Procopovici, Bucharest: Socec, 1914. (Brașov)
- CEv Pușcariu, Sextil & Procopovici, Alexie. 1914. *Carte cu învățătură*. (1581). București: Atelierele Grafice Socec & Co.
- DÎ *Documente și însemnări românești din secolul al XVI-lea (Romanian documents and notes from the 16th c)*. Gh. Chivu, M. Georgescu, M. Ioniță, Al. Mareș, Al. Roman- Moraru (eds.), Bucharest: Editura Academiei Române, 1979.
- PO Pamfil, Viorica. 1968. *Palia de la Orăștie*. 1581–1582. București: Editura Academiei.

Contact phenomena

The I-language of a bilingual

Luis López

University of Illinois at Chicago

Against the common-sense notion that bilinguals have two grammatical systems I argue that the linguistic system of a bilingual should be integrated, following ideas developed in more detail in López (2020). In particular, I argue that both the lexicon and the post-syntactic operations that lead to the externalization systems are integrated. I further argue that the distinction between code-switching and borrowing is spurious and I extend the integrated hypothesis to syntactic transfer. I use Distributed Morphology to formally describe how an integrated system may work.

Keywords: code-switching, integrated hypothesis, lexicon, gender assignment, pronominal clitics, distributed morphology, borrowing, transfer

1. Introduction

In this chapter I argue against what I call the common-sense notion of bilingualism. According to this view, the linguistic competence that is housed in the mind of a polyglot human consists of two or more discrete, countable, autonomous entities. I call this view ‘separationism’. I propose, instead, that bilinguals have a unique, integrated I-language built, like any other I-language, from internal sources in interaction with an opaque input.

I start by introducing separationism and the theoretical consequences that it has for the theory of interfaces. In this section I also briefly discuss the well-known phenomenon of Creole continua and argue that they should lead us to reject the common-sense separationist notion just described. In § 3 I present some aspects of Distributed Morphology (henceforth DM); §§ 4 and 5 I move onto presenting in sketch form two arguments for an integrated model of multilingual I-languages with only one lexicon and one PF. In §§ 6 and 7 I broaden the data base to include two other contact phenomena: ‘borrowing’ and ‘syntactic transfer’ and show how moving away from separationism can give us new insight into them.

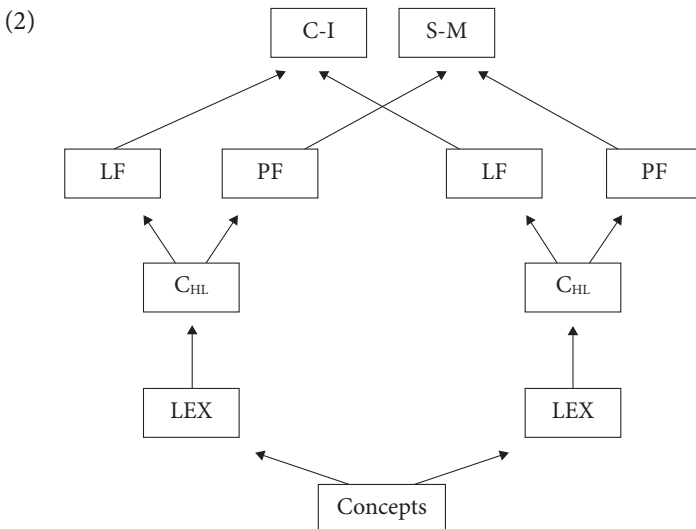
2. Separationism and architecture

As mentioned, according to a common-sense view, a bilingual has two grammatical systems in their head. We tend to view languages as countable entities and that is why we can say things like: *Mary speaks three languages*, *Joan speaks more languages than Chris does*, etc. We study languages as subjects in our school curricula, thus emphasizing their separateness. This view is also mainstream in linguistics and psycholinguistics. The quotation in (2) is an explicit statement of this approach by a famous scholar of child bilingual acquisition.

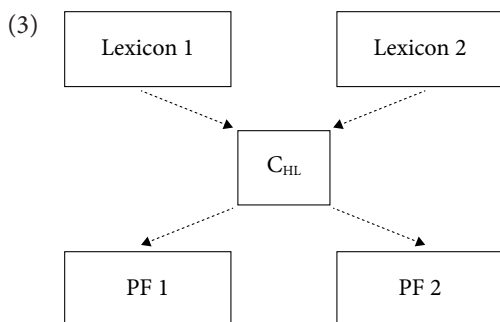
- (1) In acquiring two languages from birth, children are undergoing a sort of “double” acquisition process in which two morphosyntactic systems are acquired as fundamentally separate and closed systems. (De Houwer 2005: 43)

However, it is worth pointing out that this view is often implicit precisely because it is so commonsensical.

If indeed a bilingual has two grammatical systems, then (2) is what it looks like. Note that we should posit six interfaces for that person. The universal inventory of concepts maps onto two separate lexicons. There are two computational systems that yield two distinct PFs and LFs, which eventually map to one common cognitive-intentional system and one sensorimotor system. Nobody has explicitly claimed this picture recently, but it is implicit in the separationist assumption.



MacSwan (1999 et seq.) proposes an explicit model of bilingual grammars, and it looks like (3).



MacSwan's model incorporates Minimalist assumptions, as developed in Chomsky (1995). His contribution to our understanding of bilingualism really consists of one fundamental claim: bilinguals have two lexicons and two PFs but only one C_{HL} . Items from the two lexicons feed one computational system – and this is how code-switching comes to be. Once the computational operations within a domain are finished, the resulting structure is fed to one of the two independent PFs, one PF for each language. In later work (MacSwan & Colina 2014), PF is viewed as a system of ranked soft constraints, as in the Optimality Theory tradition. There is no exploration of LF matters in the code-switching literature and therefore I leave it aside in this chapter.

Outside linguistics proper, the notion that bilinguals have separate lexicons also has a long tradition. There is a tradition in bilingual lexical acquisition that posits that the second language lexicon is stored in a separate 'box', as in Kroll & Stewart's (1994) Revised Hierarchical Model. But I think there are good reasons to assume that bilinguals have one integrated lexicon and even one integrated PF. Let's call a model that posits two independent lexicons for bilinguals a '2Lex'.

I would now like to cast doubt on the common-sense notion that a bilingual person's I-language can be defined in terms of discrete entities. The chart in (4) represents the Creole Continuum in Guyana according to Bell (1976). This chart represents the available menu of linguistic ingredients that can go into the simple sentence *I gave him one*. These ingredients (or at least a sub-set of these ingredients) are constitutive elements of the Guayana creole speaker's I-language. These speakers can go 'up or down' this box in a manner that they feel is appropriate to the context.

(4)

1				hɪm	wʌn
2	ai				wan
3		gɛrv		ɪm	
4				i:	
5				hɪm	
6			gɪv		
7					
8	a	dɪd	gɪv	i:	
9		dɪ	gɪ		
10		dɪd			
11			gi:		
12		dɪ		hi:	
13				gɪ	
14					
15		bɪn		i:	
16	mɪ				
17			gi:	æm	
18					

Let me now invite you to try a thought experiment: imagine a linguist arrives in a territory like Guyana, call it Twin Guyana, a place where people speak like 1 and like 18 in (4) but the intermediate forms do not exist. I believe this linguist might conclude that the inhabitants of Twin Guyana are bilingual: with two lexicons and two grammars. But in real Guyana we don't find this separation; rather, as mentioned, we find that speakers can use a variety of forms. Their I-language is not neatly compartmentalized in two boxes, rather, it is a continuum of possibilities. I think Creole continua provide a stark example of why thinking of bilinguals in terms of separate systems might be misguided.

But, how about monolinguals? Could we say something like: sure, Creole speakers' I-language is a continuum, but one could not meaningfully say that a monolingual grammar is a continuum too. I disagree. As Tom Roeper and others have shown, the grammar of English includes so many nooks and crannies that, Roeper (1999) claims, English speakers all have multiple grammars in their head. I think the right way to conceptualize Roeper's claim is not as "many" grammars, but as "one" linguistic system with a range of options. We should operate under the assumption of an integrated I-language.

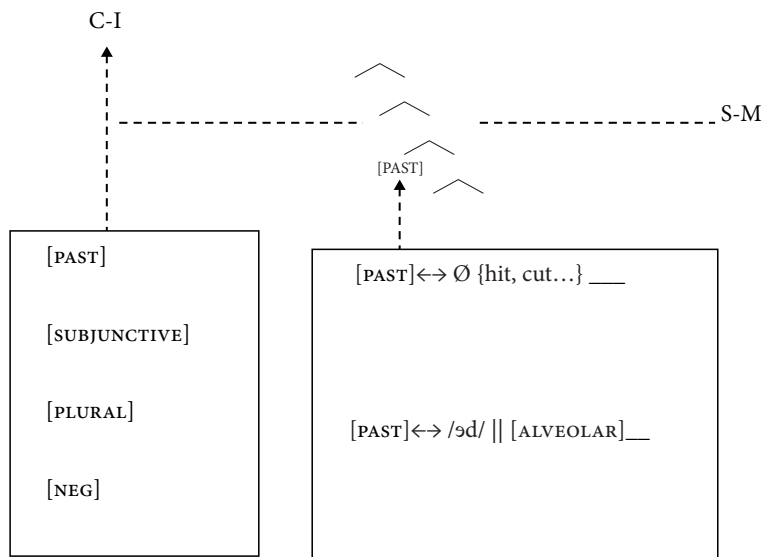
If I have convinced you that the I-language of any given person, multilingual or monolingual, is really a continuum of features, then we are ready for the next step: how do we study this continuum in a categorial manner?

3. Distributed morphology

For the analysis, I adopt Distributed Morphology (DM) (see Halle & Marantz 1993 for the seminal paper and Harley & Noyer 2000 for an introduction). The reason I adopt DM is because this framework allows us to cut the Gordian knot of linguistic continua and allows us to couch the empirical findings in an elegant solution. The model I am assuming is represented here as (5). Within DM, the traditional lexicon is now divided in two lists. List 1 includes two types of items. The first type of item is lexical roots. Following Harley (2014), I take it that the lexical roots of List 1 are nothing but abstract indices that link a syntactico-semantic representation to a phonological rule. List 1 also includes a list of grammatical features, or feature bundles, such as [plural], [negation], [past], etc.

List 2 consists of rules, called Vocabulary Insertion Rules (henceforth: VIR), which bind phonological representations to syntactic terminals. The VIRs are a stage in the process of taking the derivation toward the articulatory-perceptual systems.

(5) *Distributed Morphology framework* (see Harley & Noyer 2000, among others)



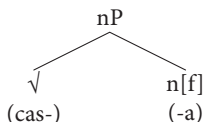
The basic syntactic structures for the clause and the noun phrase that I assume are in (6). I take it that a root may become the complement of a 'nouny' or a 'verby' categorizer, thus becoming a verb or a noun (as in Marantz 1997 et seq.). A verb becomes the complement of a Voice head, which is itself selected by T. The head of the clause is C. A noun becomes the complement of a number head, which is itself the complement of D. There are many categories not mentioned in this list:

Case, and classifiers in the noun phrase, aspect and mood in the clause, etc.; I omit a discussion of these for the purposes of this chapter.

- (6) a. Clause: C [TP T [VoiceP Voice [vP v √R]]]
 b. Noun: K [DP D [NumP Num [nP n √R]]]

In this view, roots have no grammatical information. This is an innovation of DM, with interesting consequences: grammatical information such as gender, noun class, declension or conjugation class cannot be inherent features of a lexical item. In fact, these features must be emergent, the by-product of a particular root in a particular environment. Consider now Example (7), which represents the syntactic structure of the Spanish word *casa* “house”. The word *casa* is built out of a root that spells out as *cas-* and a categorizing morpheme *n* that turns the structure into a noun. I assume throughout that gender, or noun classification generally, is a feature on *n* that combines with a sub-set of roots (see Acquaviva 2009; Kramer 2015, among many others). The feminine version of *n* spells out as /a/, in the general case.

- (7) Spanish: *casa* “house”



The same root can work as a noun, verb or adjective, as shown in the English examples below. Within DM, this means that the same root can be the complement of different categorizing heads, as shown in (8).

- (8) walk(n)
 walk(v)
 walkable(a)

The functional morphemes can also be subclassified in ‘flavors’ (Arad 2003) with more specific properties. In particular, since gender and noun class features are not inherent to the root, but emerge in a syntactic structure, it is not surprising that the same root can appear with different class features, as shown in the Swahili example in (9). In DM, we should say that the root *toto* can be the complement of various ‘flavors’ of little *n*.

- (9) (Swahili)
toto
mtoto : class 1 : child
kitoto : class 7 : small child
matoto : class 9 : big child
utoto : class 11 : childhood

4. Only one lexicon

In this section I present two empirical arguments for the notion that bilinguals have an integrated lexicon, or rather, List 1 in a DM vocabulary. The first empirical argument is based on ‘Mixed Selection’: the fact that a head from a lexicon can select a head in the other lexicon. This is impossible if bilinguals possess two separate lexicons.

Consider the examples in (10).

- (10) (Spanish)
- a. *Juan hizo la cena.*
“Juan made dinner.”
 - b. *María hizo trabajar a Santiago.*
“Maria made Santiago work.”
 - c. (Spa/Ger)
Juan hizo arbeiten.
“Juan did work” = “Juan worked” (González-Vilbazo & López 2012: 35)

In monolingual Spanish *hacer* has (at least) two uses: as a regular lexical verb of creation and as a causative verb. This is exemplified in (10a,b). In the German/Spanish code-switching (henceforth CS) variety usually referred to as *Esplugish* (González-Vilbazo 2005), *hacer* is also used as a light verb, as shown in the earlier Example (10c). The crucial datum here is the following: *hacer* can be used as a light verb only when it takes a German infinitive. If *hacer* takes a Spanish verb as a complement, the reading is causative. This is an empirical problem for a separationist thesis: if we have two independent lexicons, this restriction on *hacer* is not formulable. If we have *hacer* as part of the ‘Spanish lexicon’, it cannot select for a German verb (or a lexical item with German-like properties, this does not matter), since there are no ‘German words’ in the Spanish lexicon. If there is only one integrated List 1 this problem does not arise.

Within DM, it is very simple to provide an analysis of the mixed selection problem.

(11) [_{voiceP} Voice [_{VP} V √R]]

(12) voice ↔ /aθer/ || _____ v_{ger} *hacer* = /aθer/

Light verbs in CS are a spell out of ‘Voice’ (somewhat updating our earlier analysis). These light verbs spell out by means of a simple VIR. For concreteness, the VIR in (12) assumes a German/Spanish bilingual as in the *Esplugisch* variety where *hacer* can act as a light verb.

Rule (12) reads: “Spell out voice as *hacer* when it governs v[ger]”, the sub-index [ger] should not be understood as a ‘language label’, so thoroughly criticized in

much work on CS (MacSwan 1999). I use this index as a convenient shorthand for the cluster of grammatical properties that define the German *v* as opposed to the Spanish *v*, including the property of being able to spell out as /en/. In this model, German *v* and Spanish *v* are nothing but different flavors of *v*. Regardless of how it is formulated, the rule in (12) is only possible if there is only one lexicon or one List 1.

The second empirical argument in favor of an integrated lexicon comes from noun class or gender. I use an example of English/Swahili code-switching although I could use many other code-switching pairs. English words can be integrated into Swahili sentences quite freely. This can be seen in Example (13), where the English *certificate* appears in a Swahili sentence.

(13) (Sw/En)

Ø-saa hi-yo i-na-depend na Ø-certificate z-ako z-a Ø-shule.
 9-time dem-9 9-pres-depend with 10-certificate 10-your 10-with 10-school
 “At this time, it depends on the school certificates.”

(Myers-Scotton & Jake 2009: 339)

Swahili classifies nouns into 18 different classes – or genders – partially based on semantic criteria, partially on seemingly arbitrary grounds (see Carstens 1991, among many others). The interesting datum is that seemingly every English noun can be used in a Swahili sentence with its corresponding noun class. In Example (13), the English noun *certificate* appears to have been classified in class 9/10, as revealed by concord on the possessor *zako* and the preposition *za*. According to Myers-Scotton (2002), this assignment is permanent. And this is a puzzle for separationist theories. If we have an ‘English’ lexicon and a ‘Swahili’ lexicon, how does *certificate* end up belonging to a noun class?

In MacSwan’s sort of model, we have two options to generate a sentence like (13). Option 1 is CS proper: the noun *certificate* is an English noun, listed in the English lexicon, and it gets selected into the computational system together with some Swahili words to create a sentence like (13). But if so, where does the class feature come from? Do English words in an English lexicon have dormant class features? That does not seem very plausible. Option 2 is to take (13) to represent borrowing from one internal lexicon to the other (see § 6 for details): the noun *certificate* that appears in the sentence is really a word that has been borrowed into the Swahili lexicon. If so, it is worth pointing out that apparently every English noun can be used as in (12) (Myers-Scotton 2002). Thus, every English noun will be represented twice in the minds of Swahili bilinguals, once in each lexicon. This solution is technically feasible but probably linguistic lexicons are organized in a more efficient manner.

The solution that I propose is that the word *certificate* can appear in two different morphosyntactic frames: one, the English frame, appearing in a regular

English DP; the other, the Swahili frame, as part of a Swahili DP with class feature and concord. The analysis should, in my opinion, reflect this intuition. And the integrated model formalized via DM is precisely fitted to express it: following DM, *certificate* can be regarded as being syntactically complex, consisting of a root and a categorizing morpheme. As we know, the same root can be selected by different flavors of the same categorizing morpheme, and this is how we get *mtoto*, *kitoto*, etc. Nothing prevents the root that spells out as *certif-* from being selected by two different flavors of *n*. This is shown in Example (14):

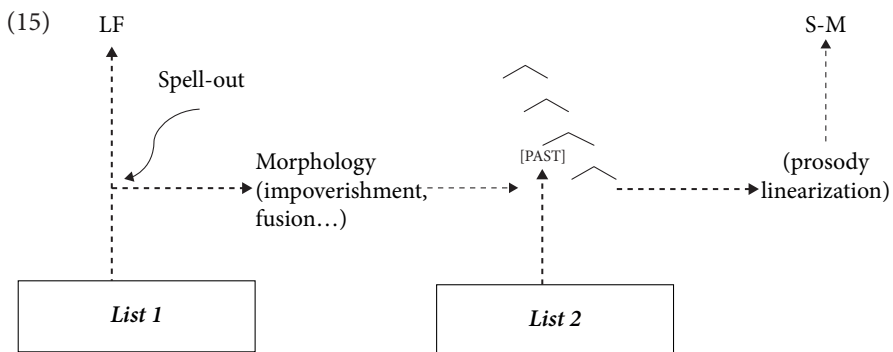
- (14) List 1
 $\sqrt{199}$
 List 2
 $\sqrt{199} \leftrightarrow /certif-/$
 $n \leftrightarrow /icate/ \parallel \text{_____} \{certif-, \dots\}$
 $n[9/10] \leftrightarrow /icate/ \parallel \text{_____} \{certif- \dots\}$

In this example I take root 199 to spell out as */certif-/* and to be the complement of two different types of *n*. The first *n* is the boring English *n* while the second *n* is the Swahili *n* that bears a noun-class marker and is able to trigger concord on the possessor and the preposition.

I submit the following conclusion to § 4: the evidence from Mixed Selection and Noun Class refutes the idea that a bilingual has two separate, autonomous lexicons. The bilingual lexicon – or List 1 in DM – is an integrated module.

5. Only one PF

In this section, I present two empirical arguments that PF is also integrated in the I-language of a bilingual. I suggest we understand the transfer of abstract syntactic structure to the sensorimotor interfaces as a sequence of operations on syntactic terminals in the DM manner (in particular see Embick & Noyer 2001). The diagram in (15) will help us with the discussion. The road from C_{HL} proper to the sensori-motor interfaces begins at ‘spell-out’. We recognize at least three stages. In the stage called Morphology there are operations that may alter the feature composition of syntactic terminals; the only important morphological operation for us is impoverishment, which deletes a subset of the features in a syntactic terminal. After operations on terminals have taken place, VIR assign phonological representations to terminals. There are still more things that need to happen to this structure before it reaches the sensorimotor interface and therefore we posit a stage three. In this stage, nuclear stress is assigned, prosodic structure is built and finally, the whole structure is linearized.



I see the whole PF as integrated in the mind of the bilingual. Here I present an empirical argument that the morphological rules are integrated as well as an argument that List 2 is integrated.

I start with morphology. The kind of evidence that I am looking for is structured as follows. Assume that in Language x there is a morphological process M_x and a language y that has a morphological process M_y . Imagine a situation in which M_x and M_y apply in exactly the same environment. Assume further that it is possible for bilingual speakers to use morphological rule M_x in a structure that is otherwise Language y and vice-versa. This scenario would suggest that Morphology is integrated among these bilinguals. In real life, it is likely that this situation is uncommon; the stars have to line up in a very precise manner. So, it is not easy to find examples. But it seems that I found one.

The clitic combination [3rd. DAT] + [3rd. ACC] is subject to rules of impoverishment in both Spanish and Catalan. The rules are somewhat different. In Spanish the dative clitic spells out as *se*, as shown in (16).

- (16) (Spanish)
Pedro le lo dijo → *Pedro se lo dijo*.
 Pedro 3.DAT 3.ACC said
 “Pedro said it to him.”

Why *se*? Spanish clitics either have a person feature or a case feature. If they have a case feature, they may inflect for plural. If they are accusative, they inflect for gender. But *se* is a clitic without person, number or case features.

- (17)
- | | | |
|---|---|----------------------------|
| <p>Cl</p> <p>person (plural)</p> <p>(<i>me, te, nos, os</i>)</p> <p><i>me</i>=1st, <i>te</i>=2nd, <i>nos</i>=1st.pl, <i>os</i>=2nd.pl</p> | <p>Cl</p> <p>case (plural)</p> <p>(<i>la(s), lo(s), le(s)</i>)</p> <p><i>la</i>=ACC.F, <i>lo</i>=ACC.M, <i>le</i>=DAT, <i>-s</i>=PL</p> | <p>Cl</p> <p><i>se</i></p> |
|---|---|----------------------------|

The impoverishment rule that eventually yields *se* for *le* can be represented in (18): the feature dative is delinked from the syntactic terminal. The resulting feature bundle can spell out as *se* but not as *le* because *le* is too rich. (this is of course inspired by the work of Halle & Marantz 1994 and Harris 1994 et seq.).

- (18) Impoverishment rule for Spanish clitic combination
[dative] → Ø / [CL _____ [accusative]]

In Catalan, the clitic combination [3rd. DAT] + [3rd. ACC] is also impoverished, but in this language the accusative clitic disappears – it is obliterated, in the sense of Arregi & Nevins (2012).

- (19) (Catalan)
El Pere li ho va dir *li ho : /liu/
The Pere 3.DAT 3.NEUT.ACC PAST say
→ *el Pere l'hi va dir.* l'hi : /li/
“Pere said it to him.”

The formula given in (20) is the rule of impoverishment for the Catalan clitic combination.

- (20) Impoverishment rule for Catalan clitic combination
[CL [accusative]] → Ø || [dative] _____

However, many bilingual speakers produce an ‘aberrant’ form in their Catalan speech, the result of applying the Spanish rule (18) and spelling it out with Catalan exponents:

- (21) (Catalan)
El Pere [s' ho] va dir
the Pere se 3.NEUT.ACC PAST say
“Pere said it to him.”

That is, in (21) the dative feature has been delinked, leaving the accusative clitic untouched. This kind of form is in fact common among bilingual speakers who are relatively free from prescriptivist rules. It can be found abundantly with a simple Google search. (The same search also finds numerous complaints by Catalan teachers, who find this sort of phenomenon deplorable.)

Forms like (22) are also found in the speech of Catalan bilinguals.

- (22) (Spanish)
Pedro le dijo.
Peter CL.DAT said
“Peter said it to him.”

One could suggest that (22) is the result of applying the Catalan impoverishment rule and spelling it out with Spanish exponents with the result that the accusative clitic disappears. However, the form shown in (22) is very common throughout the Spanish speaking world; thus, the most I can say is that (22) is consistent with my approach.

In any case, if my proposal for (21) holds water, it presents evidence that the I-language of bilinguals consists of only one Morphology module: and this includes an integrated set of impoverishment rules that may apply whenever the context is right. It is for this reason that a “Spanish” rule can apply in a Catalan context (this of course is a description of the I-language, not the performance. No doubt, there are inhibiting mechanisms that ensure that many speakers will never “cross” their morphological rules.)

Now I present one piece of evidence that the List 2 of VIRs is also integrated in one system. For that, let me take you to Gernika, Euzkadi, Spain, a town where the entire population is bilingual and code-switch frequently. One of the most frequent code-switches is between a Spanish determiner and a Basque noun and at this point one interesting puzzle arises. Spanish has gender concord which, in the general rule, spells out as [a] for feminine and [o] for masculine (but for a more complete description of gender in Spanish, see Roca 1989; Harris 1991.) When using a Spanish definite determiner, a speaker must choose between *el* (masculine) and *la* (feminine). Since Basque does not have gender concord, what determiner do these speakers use for their Basque nouns?

Badiola & Sande (2018) presented a survey to bilingual speakers from Gernika in which they were asked to choose a Spanish article for a Basque noun (the details of their methodology can be checked in the published paper.) The result of the survey is that generally they chose the masculine gender, which confirms previous arguments that masculine is default in Spanish (Roca 1989; Harris 1991). The glaring exception is shown in (23): if the word ended in [a] they consistently preferred the feminine determiner.

- (23) (Spa/Bas)
- | | | | |
|----|-----------|---------------|----------------------|
| a. | <i>la</i> | <i>makila</i> | (Spanish: el bastón) |
| | DEF.F | walking-stick | |
| b. | <i>la</i> | <i>gona</i> | (Spanish: la falda) |
| | DEF.F | skirt | |

The preference for feminine is cued exclusively to the word ending, the gender of the equivalent word in Spanish plays no role. Therefore, there is no analogical gender assignment. (By the way, this /a/ should not be confused with the Basque determiner which also spells out as [a] in Batua Euskera but spells out as [e] or [ie] in this dialect.)

The datum in (23) can be accounted for easily within an integrated theory. The VIR that links feminine with [a] is a VIR that is part and parcel of the linguistic

competence of these speakers, not a rule of the “Spanish” of these speakers. In other words, there are no two separate sets of VIRs, but just one, which will operate whenever the environment is right.

The example in (24) shows that the final /a/ in these nouns is some sort of desinence because it drops when additional suffixes are attached (and this is different from words that end in the other vowels). This confirms that the analysis is on the right track:

- (24) *hondartza*: “beach”
hondartzok: beach.ABS.DEF.PL.PROX
hondartzetan: beach.LOC.DEF

Rule (25) implements the analysis in a DM model. Basque roots like $\sqrt{\text{makil}}$, $\sqrt{\text{hondartz}}$ and $\sqrt{\text{gon}}$ can be inserted in a “Basque” morphosyntactic frame, as in (25a), with a *n* without gender or they can be inserted in an additional morphosyntactic frame with a *n* that has a gender feature.

- (25) a. *makila* : $[[\sqrt{133}] + n]$ → *makil* + *a*
 b. *makila* : $[[\sqrt{133}] + n[f]]$ → *makil* + *a*

To conclude this section, I have provided evidence that the module Morphology as well as the List 2 of VIRs are integrated in the mind of a bilingual, by which I mean that there are no rules that are specific of a language. Instead, rules are defined with respect to the environment of application and they are blind to the “language” of this environment.

6. Code-switching and borrowing

Linguists working on language contact frequently distinguish between CS and ‘borrowing’. Borrowing in this context is not understood from an E-language perspective – i.e, the fact that almost every person on earth who speaks any language has *iPad* in their vocabulary – but rather, from an I-language perspective. Roughly, the received idea is that code-switching is about the alternation of two linguistic systems in one speech act while retaining the properties of the respective systems. Instead, borrowing is taken to involve insertion of a word or idiom *W* of *L_x* into a speech act in *L_y* and full adaptation of *W* to *L_y*.

Consider the Examples (26) and (27), pronounced by English-French bilinguals in Quebec, as reported by Poplack & Dion (2012). In (26), the English word *shorts* is pronounced without a /s/ – that is, the plural feature has no exponent on the noun, as is the rule in French. This word is therefore a true borrowing (according to Poplack & Dion 2012), because the word has been fully adapted to the French grammar. In the second example, the word *tips* maintains the plural form of English: it is an example of code-switching.

(26) (Fre/Eng)

J'ai pas lavé de shorts /for/

“I haven't washed any shorts.”

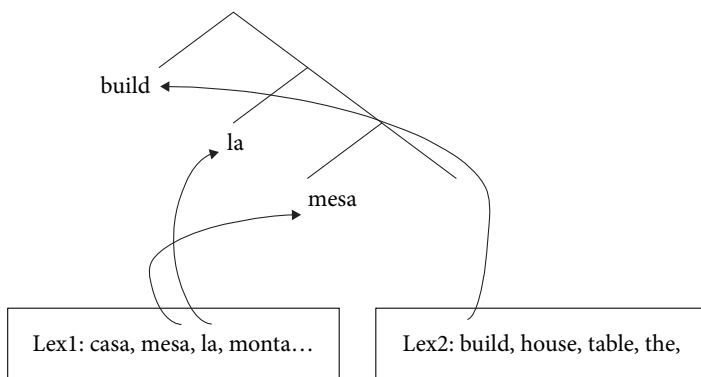
(27) *Cinq-cents piasses par trois jours, avec tes tips /tips/*

“Five hundred bucks every three days, with your tips.”

(Poplack & Dion 2012: 300–301)

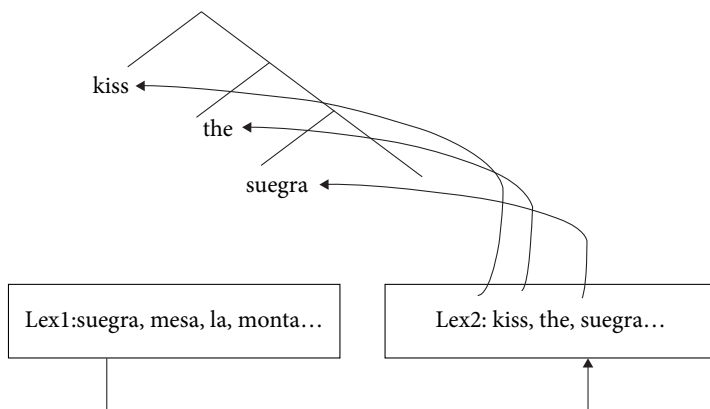
MacSwan's (1999) 2Lex theory provides us with a framework to formalize the distinction. Consider the diagrams in (28) and (29). (28) represents code-switching. In code-switching, items from both lexicons enter the derivation independently. This exemplifies *tips* in (27).

(28) Sample of a tree with code-switching. LIs from both lexicons are used to build one syntactic structure



The diagram in (29) represents borrowing. In borrowing, an item from Lexicon 1 becomes a member of Lexicon 2 with the necessary morphophonemic adjustments. This exemplifies *shorts* in (26).

(29) Sample of a tree with borrowing. LIs from Lex1 are borrowed into



It is obvious that this distinction between code-switching and borrowing can only be sustained if one assumes separationism, explicitly or implicitly. Since I don't believe bilinguals have two autonomous lexicons, this sort of analysis is not available to me. Fortunately, our framework provides the tools that we need to tackle the apparent problem.

The starting point is to understand that the I-language of an English-French bilingual has two rules that spell out a plural morpheme. One is the "English" rule that spells it out as [s] and the other is the 'French' one that spells it out as zero (and I am aware of the complications arising from liaison, etc., which I gloss over here). As usual, when referring to 'French' n or n_{fr} we are simply taking a short cut for 'flavor of n that has certain morphosyntactic properties'. The type of n decides which plural rule applies. The root 'short' is available in the "French" morphosyntactic structure and this is how /*ʃ*or/ for *shorts* is possible. On the other hand, /*t*ps/ arises because the root for *tip* can only be selected by an 'English' n .

- (30) [[[√short] n_{fr}] PL]
 [[[√tip] n_{eng}] PL]
 List 2: PL \leftrightarrow \emptyset || n_{fr} _____
 List 2: PL \leftrightarrow /s/ || n_{eng} _____

The difference between /*ʃ*or/ and /*t*ps/ arises as a consequence of the morphosyntactic framework in which they can show up. This is of course the same sort of analysis that I have been proposing above. The fact that some roots can show up in different environments is nothing new and, in particular, not a property that is circumscribed to people that we usually refer to as 'bilingual', and therefore my analysis adds nothing new to our theory of grammar. If this analysis is adopted, we conclude that the distinction between code-switching and borrowing – as a property of the bilingual I-language – turns out to be spurious.

7. Syntactic transfer

The label 'syntactic transfer', also referred to as 'calquing' or 'loan translation', entails an implicit understanding of bilingualism as involving two grammars so that one can transfer one structure to the other. Consider the example in (31):

- (31) *For this computer is this cable useful.*

This was uttered by a shop assistant in Berlin and it exemplifies a very common type of sentence heard from L2 learners of English whose L1 is German. Within separationism, it is usually described as 'transferring' the native V-2 structure onto their second language. But in fact, I suspect that there is no such thing as syntactic

transfer, as is consistent with the integrated assumption. Integrationism suggests that what we see here is that this speaker has learned new exponents or vocabulary items without any further alteration or change in his grammatical system. That is, for a particular item of List 1, he used to have just one exponent (*für*) and now he has two (*für* and *for*). Other than that, his I-language is unchanged. I believe this is a more intuitive approach to this sort of phenomenon.

Consider now (32):

- (32) (Turkish)
birin-i sorumluluk-la yükledi
 somebody-ACC responsibility-COM burden.PAST.3.SG
 “He burdened someone with the responsibility.” (Backus & Dorleijn 2009: 88)

This is a Turkish sentence uttered by a bilingual Turkish/Dutch individual. However, as Backus & Dorleijn (2009) explain, this is not a sentence that would be readily recognized by a monolingual Turkish speaker. The verb *yüklemek* (in citation form) in regular standard Turkish governs accusative on the direct object and dative on the indirect object. In this example, the indirect object receives accusative case while the direct object appears in a postpositional phrase headed by a comitative morpheme. This happens to be the case distribution of the Dutch verb *opzaleden* ‘to burden’. As Backus & Dorleijn argue, Turkish/Dutch bilinguals seem to have imposed the case distribution of the Dutch verb onto the Turkish one. But again, the integrationist framework allows us to take a second look at the phenomenon.

I propose to reanalyze (32) as follows. I again build on the idea that the same morphosyntactic structure can receive more than one spell out. In (33), the root $\sqrt{866}$ – which eventually binds to the concept ‘burden’ – merges with a verbal head that has the property of selecting an accusative beneficiary and a prepositional theme. This option can receive two spell outs in the List 2 of Turkish/Dutch bilinguals, *opzaleden* and *yüklemek*.

- (33)
-
- vP
- v $\sqrt{866}$
- [acc,PP]
- BURDEN
List 2: {opzaleden, yüklemek}

Again, there is no need to posit two separate grammars that are somehow communicated, so that one can influence the other. Adopting an integrationist standpoint simplifies our understanding of the organization of a bilingual I-language considerably.

8. Conclusions

In this chapter, I have argued for what I call an integrationist approach to the bilingual I-language. Integrationism entails a commitment to a series of independent – and independently argued – hypotheses: the multilingual individual has only one List 1 of roots and functional categories, understood as abstract features. She has also one Morphology and one List 2 of vocabulary items. The empirical evidence presented here is based on the phenomena of Mixed Selection in German/Spanish code-switching, noun class incorporation in Swahili/English code-switching, clitic combinations among Catalan/Spanish bilinguals and gender in Basque/Spanish code-switching.

The integrationist assumptions shed new light not only on code-switching but also on borrowing and syntactic transfer and lead to the conclusion that all these apparently disparate contact phenomena have a common denominator. This commonality is highlighted with an analysis based on DM.

As a positive consequence of this simplification, we can safely conclude that the multilingual individual has a linguistic module with one interface to the conceptual-intensional module and one interface with the sensorimotor modules and not a separate pair of interfaces for each language.

Before closing, I would like to point out that these proposals were not born in a vacuum. The separationist view has recently been questioned in several quarters. Goldrick, Putnam & Schwartz (2016) proposed an integrated model based on soft constraints (see López 2016 for discussion). Pierantozzi (2012); Bandi-Rao & Den Dikken (2014); Grimstad et al. (2014); Alexiadou et al. (2015), and Riksem (2017) have also provided analyses of code-switching phenomena through a DM prism. Beyond the immediate horizon, the co-activation facts have led many psycholinguists to conclude that “all the languages that are known and used become part of the same language system” (Kroll et al. 2015: 377). Finally, many scholars in sociolinguistics and education are now engaged in a movement called ‘translanguaging’, according to which “the mental grammars of bilinguals are structured but unitary collections of features, and the practices of bilinguals are acts of feature selection, not of grammar switch” (Otheguy, García & Reid 2015: 281). I could not agree more – although, I would like to point out, empirical argumentation like the one presented in this chapter is necessary for any new approach.

References

- Acquaviva, Paolo. 2009. Roots and lexicality in Distributed Morphology. *York Papers in Linguistics Series 2*. 1–21.
- Alexiadou, Artemis, Terje Lohndal, Tor Afarli & Maren Berg Grimstad. 2015. Language mixing: A Distributed Morphology approach. In Thuy Bui & Denis Ozildiz (eds.), *Proceedings of NELS*, 25–38. Amherst, MA: GLSA Publications.
- Arad, Maya. 2003. Locality constraints on the interpretation of roots: The case of Hebrew denominal verbs. *Natural Language and Linguistic Theory* 21. 737–778.
<https://doi.org/10.1023/A:1025533719905>
- Arregi, Karlos & Andrew Nevins. 2012. *Morphotactics*. Studies in Natural Language and Linguistic Theory 86. Dordrecht: Springer. <https://doi.org/10.1007/978-94-007-3889-8>
- Backus, Ad & Margreet Dorleijn. 2009. Loan translations versus code-switching. In Barbara Bullock & Jacqueline Toribio (eds.), *The Cambridge handbook of linguistic code-switching*, 75–94. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511576331.006>
- Badiola, Lucía & Ariane Sande. 2018. Gender assignment in Basque mixed Determiner Phrases: A study of Gernika Basque. In Luis López (ed.), *Code-switching – experimental answers to theoretical questions*, 15–38. Amsterdam: John Benjamins. <https://doi.org/10.1075/iHLL.19.02bad>
- Bandi-Rao, Shoba & Marcel Den Dikken. 2014. Light switches: On v as a pivot in code-switching, and the nature of the ban on word-internal switches. In Jeff MacSwan (ed.), *Grammatical theory and bilingual code-switching*, 161–184. Cambridge, MA: MIT Press
- Bell, Roger T. 1976. *Sociolinguistics: Goals, approaches and problems*. London: Batsford.
- Carstens, Vicki. 1991. DP Structure in KiSwahili. PhD Dissertation: UCLA.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- De Houwer, Anneke. 2005. Early bilingual acquisition. In Judith Kroll & Annette de Groot (eds.), *Handbook of bilingualism*, 30–48. Oxford: Oxford University Press.
- Embick, David & Rolph Noyer. 2001. Movement operations after syntax. *Linguistic Inquiry* 32 (4). 555–595. <https://doi.org/10.1162/002438901753373005>
- Goldrick, Matt, Michael Putnam & Linda Schwartz. 2016. Coactivation in bilingual grammars: A computational account of code mixing. *Bilingualism: Language and Cognition* 19. 857–876. <https://doi.org/10.1017/S1366728915000802>
- González-Vilbazo, Kay. 2005. Die syntax des code-switching. Ph.D. Dissertation: Universität zu Köln.
- González-Vilbazo, Kay & Luis López. 2012. Little v and parametric variation. *Natural Language & Linguistic Theory* 30(1). 33–77. <https://doi.org/10.1007/s11049-011-9141-5>
- Grimstad, Maren Berg, Terje Lohndal & Tor A. Åfarli. 2014. Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian. *Nordlyd* 41(2). 213–237. <http://septentrio.uit.no/index.php/nordlyd>
- Halle, Morris & Alec Marantz. 1993. Distributed morphology and the pieces of inflection. In Ken Hale & S. Jay Keyser (eds.), *The view from building 20*, 111–176. Cambridge, MA: MIT Press.
- Halle, Morris & Alec Marantz. 1994. Some key features of Distributed Morphology. *MIT Working Papers in Linguistics* 21. 275–288.
- Harley, Heidi. 2014. On the identity of roots. *Theoretical Linguistics* 40(3/4). 225–276. <https://doi.org/10.1515/tl-2014-0010>
- Harley, Heidi & Rolph Noyer. 2000. Distributed Morphology. *Glott International* 4(4). 3–9.

- Harris, James. 1994. The syntax-phonology mapping in Catalan and Spanish clitics. *MITWPL* 21. 321–353
- Harris, Jim. 1991. The exponence of gender in Spanish. *Linguistic Inquiry* 22. 27–62.
- Kramer, Ruth. 2015. *The morphosyntax of gender*. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199679935.001.0001>
- Kroll, Judith F. & Erika Stewart. 1994. Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations. *Journal of Memory and Language* 33. 149–174. <https://doi.org/10.1006/jmla.1994.1008>
- Kroll, Judith F., Paola E. Dussias, Kinsey Bice & Lauren Perrotti. 2015. Bilingualism, mind, and brain. *Annual Review of Linguistics* 1. 377–394.
<https://doi.org/10.1146/annurev-linguist-030514-124937>
- López, Luis. 2016. Questions on data and input to GEN. *Bilingualism, Language and Cognition* 19. 889–890. <https://doi.org/10.1017/S1366728916000079>
- López, Luis. 2020. *Bilingual grammar: Toward an integrated model*. Oxford: Oxford University Press. <https://doi.org/10.1017/9781108756181>
- MacSwan, Jeff. 1999. *A minimalist approach to intrasentential code switching*. New York: Garland Press.
- MacSwan, Jeff & Sonia Colina. 2014. Some consequences of language design: Codeswitching and the PF interface. In Jeff MacSwan (ed.), *Grammatical Theory and Bilingual Codeswitching*, 185–210. Cambridge, MA: MIT Press
- Marantz, Alec. 1997. No escape from syntax: Don't try morphological analysis in the privacy of your own lexicon. In Alexis Dimitradis, Laura Siegel, Clarissa Surek-Clark & Alexander Williams (eds.), *Proceedings of the 21st Penn Linguistics Colloquium*, 221–225. University of Pennsylvania Working Papers in Linguistics, Philadelphia: University of Pennsylvania.
- Myers-Scotton, Carol. 2002. *Contact linguistics*. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780198299530.001.0001>
- Myers-Scotton, Carol & Janice Jake. 2009. A universal model of code-switching and bilingual language processing and production. In Barbara Bullock & Jacqueline Toribio (eds.), *The Cambridge handbook of linguistic code-switching*, 336–357. Cambridge: Cambridge University Press.
- Otheguy, Ricardo, Ofelia García & Wallis Reid. 2015. Clarifying translanguaging and deconstructing named languages: A perspective from linguistics. *Applied Linguistics Review* 6. 281–307.
<https://doi.org/10.1515/applirev-2015-0014>
- Pierantozzi, Cristina. 2012. Agreement within early mixed DP. In Kurt Braunmüller & Christoph Gabriel (eds), *Multilingual individuals and multilingual societies*, 137–152. Amsterdam: John Benjamins. <https://doi.org/10.1075/hsm.13.10pie>
- Poplack, Shana & Nathalie Dion. 2012. Myths and facts about loanword development. *Language Variation and Change* 24. 279–315. <https://doi.org/10.1017/S095439451200018X>
- Riksem, Brita. 2017. Language mixing in American Norwegian noun phrases. Ph.D. Dissertation: Norwegian University for Science and Technology.
- Roca, Iggy. 1989. The organisation of grammatical gender. *Transactions of the Philological Society* 87(1). 1–32. <https://doi.org/10.1111/j.1467-968X.1989.tb00617.x>
- Roeper, Tom. 1999. Universal bilingualism. *Bilingualism: Language and Cognition* 2: 169–186.
<https://doi.org/10.1017/S1366728999000310>

-*η* plurals in North Lombard varieties

Differential plural marking and phases

M. Rita Manzini, Leonardo M. Savoia and Benedetta Baldi

Università di Firenze

We focus on North Lombard *-a* and *-n* feminine plurals, for which we provide a morphological analysis. At the syntactic level, the relevant varieties are characterized by the phenomenon of Differential Plural Marking, whereby phasal domains have different realizations of plural morphology on the head of the phase and on the complement of the phase. We provide an account of DPM based on the assumption that under the PIC the complement of the phase and its head are externalized separately. We draw consequences concerning clitics as phasal heads as well as object agreement with participles and with finite verbs.

Keywords: number, nominal class, agreement, phase, clitic

1. Differential plural marking in Mesolcina Valley varieties

The main topic investigated in this work is what we call Differential Plural Marking (DPM), i.e., the fact that plural inflections are selectively externalized on the constituents that could in principle bear them. Because of space limitations, we illustrate only varieties of the Mesolcina Valley,¹ based on data from Manzini & Savoia (2005) from Soazza speakers. The distribution of plural feminine *-əη* includes Ns, but not articles, as shown in (1a). In (1b–c–d) *-əη* is also lexicalized on demonstratives and the other pre-nominal adjectives and quantifiers, as well as in post-nominal adjectives, in (1e). The singular has *-a*, as in (1a'–b'–c'–d'). Masculine plural *-i* occurs on articles and other determiners such as demonstratives (2a–b), but not in other pre-nominal modifiers such as quantifiers (2c). Moreover *-i* characterizes a subset of masculine Ns and adjectives whose base ends in *-l*, for instance (2a).

1. The Mesolcina Valley is an alpine valley of Italian Switzerland, located in the Canton Grisons except for some villages located in Canton Ticino.

Another plural morpheme *-η* occurs in a small set of human nouns as in (2b).² The masculine singular is essentially uninflected, as in (2a'–b').

- (1) a. *l-a fkaβel-əη*
 the-F chair-PL
 “the chairs”
 a'. *l-a fkaβel-a*
 the-F chair-F
 “the chair”
 b. *kwel-əη/pək-əη fkaβel-əη*
 that-FPL/few-FPL chair-FPL
 “those / few chairs”
 b'. *kwel-a fkaβel-a*
 that-F chair-F
 “that chair”
 c. *l altr-əη fkaβel-əη*
 the other-FPL chairs-FPL
 “the other chairs”
 c'. *l altr-a fkaβəl-a*
 the other-F chair-F
 “the other chair”
 d. *l-a mε-η sorel-əη*
 the-FP my-FPL sister-FPL
 “my sisters”
 d'. *l-a mε sorel-a*
 the-F my sister-F
 “my sister”
 e. *do botiλλ-əη pjən-əη*
 two bottle-FPL full-FPL
 “two full bottles”

2. The plural *-i* in masculine nouns ending in *-l*, like *martel* “hammer” in (2a), derives from an original sequence *-ll-i* which has palatalized to *-i*. So, we can speak of a morpho-phonologically specified subset of nouns alternating *-l* in the singular and *-i* in the plural. The *-η* plural in masculines and feminines like *mat/ma'toη* “boy/boys”, *mata-ma'taη* “girl/ girls”, *femna-fem'naη* “woman/ women” does not seem to be connected with the feminine *-əη* analyzed in this chapter. Indeed, it does not interact with the verbal 3rd plural inflection as the feminine *-əη* does, it has a limited distribution in a subset of nouns and, differently from feminine *-η*, it attracts the stress on the preceding nucleus (Salvioni 1902; Rohlf's 1968[1949]), implying also a different etymology. In other words, this morphology is lexically determined. The alternation *əm/ oməη* “man/men” in (2b) is to be traced back to etymological syllabic difference between *homo/homines* “man/ men” in Latin, preserved in many Italian varieties, including standard Italian *uomo/uomini*. This allomorphy implies two specialized lexical bases, for the singular and for the plural.

- (2) a. *i mar'te-i*
 the.MPL hammer-MPL
 “the hammers”
- a'. *el mar'tel*
 the hammer
 “the.M hammer”
- b. *kw-i om-əŋ/ma't-oŋ/di:t*
 that-MPL man-PL/boy-PL/finger
 “those men/boys/fingers”
- b'. *kwel om/mat/di:t*
 that.M man/boy/finger
 “that man/boy/finger”
- c. *pɔk oməŋ*
 few man.PL
 “few men”

The partial distribution of plurals observed in the Mesolcina varieties, most clearly in the feminine (1), is far from isolated. Costa & Figureido Silva (2002) show that in Brazilian Portuguese dialects, plural *-s* occurs on (masculine and feminine) Ds, but not on Ns and predicative As – mirroring the North Lombard distribution. In Ladin (Rasom 2008), plural *-s* is again involved, but the distribution is the same as in Lombard plurals, namely on Ns and As, to the exclusion of Ds. Additional examples from North Italian and Rhaeto-Romance (Ladin, Friulian) varieties are provided by Savoia, Baldi & Manzini (2018); Baldi & Franco (2018); Manzini, Baldi & Savoia (2018).

The opposition between D and N/A cannot be lexical or morphological for more than one reason. First, there are no morphophonological reasons why the bases that support *-a* inflections should not also support *-η* inflections or vice versa. Second, and more important, the same categories may have one or the other inflection according to syntactic position. Thus, as highlighted by the literature just quoted, in many varieties (e.g., Ladin, Friulian) prenominal As in the functional area of the DP inflect like Ds, while postnominal As inflect like Ns.

Now, generally in North Lombard varieties, the nominal plural *-η* for the feminine is identical to the finite verb plural *-η* (and according to external accounts borrowed from it). Subject clitics in (3) take the form that one might expect from inspecting the determiners in (1)–(2), namely *əl/i* for the masculine singular and plural respectively and *la* for both the feminine singular and the feminine plural. With feminine plural subject clitics the *-əŋ* plural inflection is present on the finite verb. In contrast, it is absent with masculine subject clitics, although it seems to be marginally acceptable.

- (3) a. *ə/l-l-a dɔrm*
 3.M/3-F sleep
 “S/he sleeps”
- b. *i dɔrm^(*/?) -əŋ*/ *l-a dɔrm-əŋ*
 3.MPL sleep/ 3-F sleep-FPL
 “They sleep”

Consider next object clitics; these again take the form *l/i* for the masculine singular and plural respectively and *la* for both the feminine singular and the feminine plural. In the presence of feminine plural OClS, *-əŋ* again adds to the finite verb, as in (4b–b’). In fact, in (4b’) *-əŋ* can be seen to add to the 1st person singular subject inflection *-i*. This kind of object agreement with the finite verb is exceedingly rare or otherwise inexistent in Romance. It seems directly connected to the capability of *-əŋ* to function as a finite verb inflection as well as a nominal (adjectival, participial, etc.) inflection.

- (4) a. *tu l/l-a/i ve:t*
 you 3.M/3-F/3.MPL see
 “You see him/her/them(m.)”
- b. *tu l-a ved-əŋ*
 you 3-F see-FPL
 “You see them(f.)”
- b’. *l-a tʃam- i- əŋ*
 3-F call- 1SG- FPL
 “I call them(f.)”

As we may expect, perfect participle displays *-əŋ* agreement with feminine OClS understood as plural (5a). Masculine perfect participles (5b) can also be endowed with the agreeing plural morphology *-i*, like some of the Ns in (2). Nevertheless, other perfect participles or predicative adjectives have zero inflection in the masculine plural (5d) – as opposed to the *-əŋ* inflection of the feminine plural (5c).

- (5) a. *tu me l a -i -ŋ portad-əŋ*
 You me 3 have-2SG-FPL brought-FPL
 “You have brought them(f) to me”
- b. *i a'veŋ lava-i*
 3.MPL we.have washed-MPL
 “We have washed them(m.)”
- c. *l-a ε-ŋ tʃrak-əŋ/vənitʃ-əŋ ier*
 3-F be-FPL tired-FPL/come-FPL yesterday
 “They(f.) are tired/have arrived yesterday”

- d. *i ε strak/vənitf ier*
 3.MPL be tired/come yesterday
 “They(m.) are tired/have arrived yesterday”

Finally, note that object clitics may occur in enclisis, for instance in the infinitive in (6a) and in imperative in (6b–c), presumably because of lexicalization of the imperative verb in C. In this instance, we observe the orders in (6), where the inflected imperative is followed by the *l-* clitic and then by the *-əη* plural inflection. Perhaps the simplest way to construe these data is as an instance of mesoclis. In the analysis of Manzini & Savoia (2011, with references), the lexicalization of the infinitive and the imperative in C strands the object inflection in the I domain, and the enclitic occurs between them.

- (6) a. *l ε dre a kuzi-l-əη*
 3.SG is behind to sew-3-FPL
 “(S)he is sewing them(f.)”
 a. *tfam- a- l- əη*
 call- 2SG- 3- FPL
 “Call them(f.)!”
 b. *tfama- də- l- əη*
 call- 2PL- 3- FPL
 “Call them(f.)!”

A consequence of the distribution of agreement described so far is that examples of Soazza involving both a subject and an object feminine clitic give rise to multiple ambiguities in case of plural agreement with the verb. Consider (7a); the *-əη* morphology on the verb could refer either to the subject clitic or to the object clitic. In fact, matters are even more complex, since a single *-əη* can imply the plurality of both subject and object, so that (7a) is three-way ambiguous. In (7b) the ambiguity is reduced by the presence of a perfect participle bearing an *-əη* agreement which can only be referred to the object. Therefore (7b) is two-ways ambiguous. Finally, in the contexts with a feminine plural lexical object, as in (7c), *-əη* occurs as the plural inflection of the noun to the exclusion of the verb. In short, the occurrence of *-əη* is associated to the plural interpretation of the subject or the object clitic (see § 3.1).

- (7) a. *l-a l-a tfam-əη*
 3-F 3-F call-FPL
 “She calls them(f.)/They(f.) call her/They(f.) call them(f.)”
 b. *l-a l a-η tfamad-əη*
 3-F 3 have-FPL called-FPL
 “She has called them(f.)/They(f.) have called them(f.)”

- c. *kwel-a femn-a l-a lav-a l-a kamiz-əŋ*
 that-F woman-F 3-F wash-3P the-F shirt-FPL
 “That woman washes the shirts”

The distribution of plural morphology in the DP, in the vP (i.e., with participles) and in the CP (i.e., with finite verbs), both in the feminine and in the masculine, is as summarized in (8). Determiners and pronominal clitics have a morphology which contrasts with that of lexical categories, whether nominal or verbal.

- (8) a. DP: D A N A
 F *-a* || *-əŋ* *-əŋ* *-əŋ*
 M *-i* || *-i/-ŋ/∅*
 b. CP: SCI I
 F *-a* || *-əŋ*
 M *-i* ||
 c. vP: OCl V A
 F *-a* || *-əŋ* *-əŋ*
 M *-i* || *-i/∅* ∅

The left-right asymmetry is systematic. However, in the masculine the only elements that obligatorily bear the plural morphology *-i* are determiners and clitics, while *-i* only occasionally surfaces on Ns and As. In the feminine, the distribution is apparently reversed, since Ns, As and Vs present the plural morphology *-əŋ*, while determiners and clitics lack it.

2. Morphological analysis: *-a* as a plural

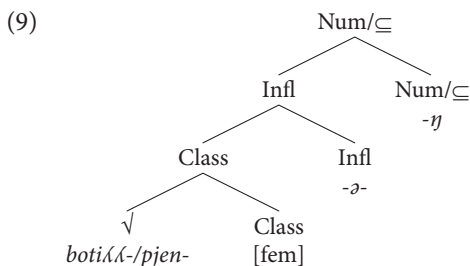
As summarized at the end of the previous section, the Soazza language in (8) descriptively presents a double asymmetry in the plural. One asymmetry is syntactic in nature and holds between Ds and lexical categories in the DP and in the sentence. This will be discussed in the next section. In this section, we discuss the morphological asymmetry between feminine and masculine. The feminine/masculine asymmetry is unexpected if we consider related phenomena in other Romance languages. Thus the partial distribution of plural *-s* in Brazilian Portuguese dialects (Costa & Figueredo Silva 2002) equally affects the *-o* masculine class and the *-a* feminine class alike. The masculine/feminine asymmetry is not isolated either, since it equally characterizes Ladin (Rasom 2008).

We assume a model of the internal organization of nominal categories based on the idea that morphemes are combined by the syntactic operation of Merge, along the lines of Distributed Morphology (DM). We differ from the latter in assuming

that the input of Merge is formed by classical lexical entries endowed with both a phonological and a semantic content. In other words, all lexical entries, including inflectional material, are endowed with interpreted content governing their occurrence in the sentence. Along the same lines, we exclude the recourse to morphological rules able to manipulate the interpretive properties associated to lexical entries. Thus we assume projection from the lexicon, and reject Late Insertion in the spirit of what seems to us the true Chomskyan minimalist approach.

In Romance, the lowest component of a Noun is a root (or derived lexical base) with predicative content (Higginbotham 1985), for instance *libr-* in Italian or Spanish *libr-o* “book”. This combines with nominal class specifications including gender (e.g., masculine, feminine), which restrict the argument of the predicative base. A vocalic morpheme encodes properties that include both nominal class and declension class, for instance *-o* in *libr-o*. This is sometimes referred to as Th (thematic vowel); here we label it Infl. Finally, an additional slot may be available, specialized for number (e.g., Spanish) or in other Indo-European languages for case (e.g., Albanian). This is the slot where the sigmatic plural of Western Romance is realized, e.g., Spanish *libro-s* “books”. As discussed on several occasions (Manzini et al. 2018; Manzini & Savoia 2018b; Savoia, Baldi & Manzini 2019), selectional restrictions of the nominal roots on gender and on the inflections with which the root combines, are able to account for their distribution.

Based on the above discussion, we propose the structure in (9) for *botiλλ(ə)η/pjenəη* “bottles/full(fpl.)” in (1b). As already mentioned, current approaches follow the idea that gender is a property merged with the root whereas number (plural) is introduced by means of a higher inflection. This split is represented in (9), where the predicative root is combined with nominal class specifications, which include gender, while *-η* lexicalizes the specialized category of plural. This is the same analysis applied by the literature to Spanish *-s* and extendable to other specialized plural morphemes. We suggest the content of plurality by the subset label \subseteq , consistent with the conceptualization of plural as attributing divisibility to the denotatum of the root. The *-η* plural in (9) is then externalized in the specialized Num/ \subseteq node.



A different situation with respect to (9) is shown by languages like Italian, where plural implies a change in Infl morphemes. Manzini & Savoia (2018a, 2019) argue that plural can be represented together with gender in Class, and hence be realized by Infl. Thus in Italian *-o* alternates with *-i* in *libro/libri* “book/books” under Infl. The representation of plural in Class is consistent with the conceptualization of plural as attributing divisibility to the denotatum of the root, i.e., a classifier property not substantially different from gender (see also Déchaine et al. 2014).

The conclusion that there must be at least two different loci for plurality, a lower and a higher one is independently attested in the literature. Thus the Italian morpheme *-a* acts as the feminine singular but also as the plural in a subset of nouns. Acquaviva (2008: 123) identifies the *-a* plurals in Italian as lexical plurals, i.e., plurals obtained by means of morphological mechanisms different from the ones that generate regular plurals, and assumes that they are to be treated as “distinct, inherently plural nouns, related to the base noun by a word-formation process”. In other words, they would not be a bona fide plural inflection, but an exponent for a different lexeme. A solution substantially derived from Acquaviva (2008) is also discussed in Alexiadou (2011) with respect to mass noun plural in Greek, distinguishing between the grammatical plural and the lexical plural, associated to the nominalizing head *n* of DM.

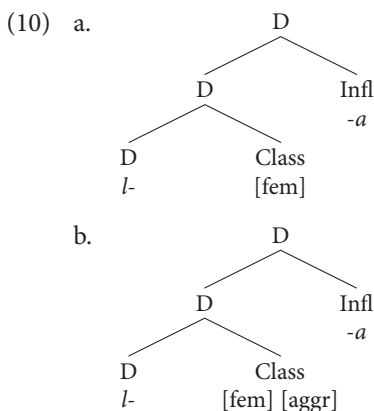
Importantly, the existence of two plurals, one structurally closer to the root and the other higher in the structure is consistent with a purely inflectional treatment of plurals. Kramer (2016), in dealing with the number system of Amharic, distinguishes regular (concatenative) plurals from irregular (non-concatenative) ones, lexically restricted to special set of nouns. She considers regular plurals as the realization of Num. Lexically restricted plurals are determined by an inherent property of the head also hosting gender, namely *n* in DM (Class here).

Against this background, let us then consider the determiner *l-a* of Soazza, surfacing in the feminine singular and plural. Two different morphosyntactic analyses of the *-a* inflection are possible. The traditional analysis, adopted by formal treatments like Rasom (2008) and Bonet (2018) for Ladin, is that *-a* is the feminine default, surfacing in the absence of plural *-η/-s*. The view that *-a* is a default yields a description of the facts in (8), whereby no plural is externalized on feminine Ds, while masculine Ds obligatorily realize number. In other words, masculine and feminine plurals present opposite distributions.

An alternative account is suggested by the independent existence of *-a* feminine plurals in standard Italian, denoting either collective and mass plurals or count plurals of ‘weakly differentiated’ individuals (Acquaviva 2008), e.g., *uov-a* “eggs”. Manzini & Savoia (2018a, 2019) reject the lexical treatment of Acquaviva, alluded to above, in favor of an inflectional treatment. Specifically, they characterize the

number property externalized by *-a* as [aggregate], i.e., in terms of the notion that Chierchia (2010) uses to characterize the common core of mass and plural denotation – which they associated with the Class node.

Returning to Soazza, then, we propose that the *-a* inflectional morphology is not a feminine default. Rather, it externalizes not only the [fem] nominal class, as in (10a), but also the [aggr] property which, as defined above, is compatible with plural denotation, as in (10b). In (10b) nominal class specifications include gender and number, as suggested by Manzini & Savoia (2018a, 2019) for vocalic plurals in general.



The D in (10b) enters Agree with Ns/As in (9), not in that *-a* realizes the [fem] default, but in that it realizes the specification [aggr] which is non-distinct from [\subseteq]. To be more precise, we assume that set-divisibility [\subseteq] is a specialization of [aggr] (which also encompasses mass denotation), so that the two are compatible under Agree. As for the *-ŋ* plural morphology of verbs, we assume that it shares the same content as the nominal inflection, namely [\subseteq].

As pointed out in § 1, the asymmetric lexicalization of the plural between D and N inside DPs is well-documented in Italian dialects (Savoia et al. 2018, 2019; Manzini et al. 2018) as well as in other Romance languages (cf. Costa & Figueiredo 2002 on Brazilian Portuguese). It may be expected that denotational properties of D and N are reflected by different plural inflections and that D can select a specialized inflection, if available in the lexicon. So, this distribution may be ultimately imputed to the role determiners play in the referential anchoring of arguments. However in the discussion that follows we propose that this result is mediated by the computational notion of phase.

3. Syntax: Externalization by phase

We turn now to syntax and to what we have called Differential Plural Marking (DPM), as manifested by left-right asymmetries in the DP and in the sentence summarized in (8). Accounts previously proposed in the formal literature restrict themselves to the particular form that DPM takes in a given language and are difficult to extend to a crosslinguistic account. This is true of Costa & Figueiredo's (2002) DM approach to Brazilian Portuguese, and Rasom's (2008) cartographic approach to Ladin (though see § 3.2). By contrast, we are beginning to form a picture of DPM along the lines of (11), where the syntactic distribution of DPM is governed by the universal structural oppositions in (11a), while the morphological externalization of these structural oppositions varies in an essentially arbitrary fashion, as in (11b).

- (11) DPM Generalization
- a. *Syntactic universals*
D vs. NP, OCl vs vP, ...
 - b. *Externalization parameter*
plural externalized
 - i. by -s
by -i
by -n
by \emptyset , ...
 - ii. by the same morphology
throughout
by different morphology on D
vs. NP

Our main proposal concerns the syntactic universals underlying the DPM Generalization. Focussing on the DP-internal distribution of plural morphology, we propose that the D vs NP asymmetry corresponds to the distinction between head of a phase and complement of the phase, independently individuated by the Phase Impenetrability Condition (PIC) in (12). We also assume the inheritance mechanism whereby particular ϕ -features are inherited by T from the phase head C (Chomsky 2008).

(12) *Phase Impenetrability Condition*

In a structure $[_{ZP} Z...[_{HP} \alpha [_H YP]]]$, where Z and H are phase heads the domain of H is not accessible to operations at ZP; only H and its edge are accessible to such operations. (PIC, Chomsky 2001: 13, 14)

Though the original phases proposed by Chomsky (2001) include only vP and CP, there are reasons to conclude that DP is a phase, beginning with the evidence

that led it to be considered as a cyclic node for Subjacency, namely the difficulty in extracting from it. We therefore assume CP, vP and DP to be phases. We also assume that the operation Agree is at play within the DP as within the sentence. It is sometimes concluded that Concord within the DP must be a different operation with respect to sentential Agree (Rasom 2008 on Ladin is an example of this). However there are many proposals in the literature as to how to extend Agree to DPs, including Multiple Agree (Carstens 2000).

Thus the universal core of the DPM generalization in (11a) consists of the PIC and of the operation Agree, as in (13).

- (13) *DPM: Syntactic universals*
- i. PIC
 - ii. Agree

As for the DPM parameter in (11b)(ii), nothing much needs to be said about languages where plural is consistently realized on all nominal categories. The problematic varieties are those like Soazza where externalization of plural is sensitive to the boundary between head of the phase (namely D and D-like categories) and complement of the phase (roughly NP, namely N and AP modifiers of N). Now, the notion of phase and the PIC are tied up with that of Transfer of syntactic objects to the interfaces, hence with externalization (Spell-Out):

syntactic objects ... at some point in the derivation, are transferred to the two interfaces: transfer to the sound interface is often called “Spell-Out.” ... We therefore hope to be able to establish a “Phase Impenetrability Condition,” which guarantees that mappings to the two interfaces can forget about what they have already done, a substantial saving in memory. (Chomsky 2005: 16–17)

We can characterize varieties like Soazza as enforcing a uniform lexicalization of φ -features at each Transfer point. This implies a uniform lexicalization of plural within the NP complement of the DP phase. Yet it leaves the possibility open that at the immediately successive transfer point, involving the D head of the phase, a lexicalization different from that of NP may be chosen.

Thus under (14) languages may enforce a uniform lexicalization of φ -features all the way (14i). They may also enforce uniformity within each Spell-Out domain (14ii), which under the PIC is identified as the complement of a phase. In this latter instance no uniformity is predicted with other Spell-Out domains, hence specifically with the head of the phase.

- (14) *DPM: Externalization parameter*
- Occurrences of the φ -feature (set) F, are externalized
- i. uniformly
 - ii. uniformly in each Spell-Out domain ...

The adoption of value (ii) of the parameter in languages like Soazza means in particular that, as schematized in (8), the feminine plural is uniformly lexicalized by *-əŋ* in the complement of the D head, namely NP (N and As) whereas a different lexicalization *-a* characterizes the phase head D. For the masculine, it is necessary to keep in mind that the plural morphology on N is lexically determined; in the absence of lexically determined plurals, the productive patterns involve no plural morphology. By contrast, the D head of phase presents the obligatory externalization of *-i*.

3.1 The vP and CP phases

The account of DPM in (13)–(14) predicts that we ought to be able to find reflexes of it in the sentential phases, namely vP and CP. If we push an analysis along the lines of (13)–(14), then interesting syntactic consequences follow, which will be noted here. Let us begin with the vP phase, as summarized in (8c), focusing on the feminine data, which are more clear-cut. The fact that OCl=D and that participial V=A suggests, everything else equal, that OCl is the head of the vP phase, while the participle, together with its complements and modifiers, i.e., VP, is the complement of the phase. This corresponds to a representation of the type in (15) for Example (5a) *l(a) ... portadəŋ* ‘brought them’.

(15) ..._[v/AccVoice] *la* _[VP] *portad-əŋ* ... cf. (5a)

The structure in (15) effectively says that the accusative clitic is an exponent of the *v* head responsible for the accusative alignment of the sentence (and possibly for the merger of the external argument, unless the latter is made to depend on Voice). This conclusion in turn is not isolated in the literature. The mainstream analysis of Romance clitics is Kayne’s (1991), whereby clitics are first-merged in argument position and subsequently adjoined to some functional projection on the sentential spine. The structure in (15) corresponds to the leading formal alternative proposed by Sportiche (1996), namely that clitics are first-merged in a functional position along the sentential spine, projecting AccVoice (for the accusative clitic) and other Voice positions. Another way to phrase this is that object clitics are the agreement head of the vP phase, AgrO.³ This conclusion is substantiated by the fact that the presence of a feminine plural lexical object does not select the *-əŋ* inflection on

3. Roberts (2010: 57) proposes that ‘the label of (active, transitive) *v** contains ... unvalued versions of the very ϕ -features that make up the clitic’. Hence ‘the clitic’s label is not distinct from *v**’s. Thus the clitic can adjoin to *v** and form a derived minimal head’, as in (i). His aim is to circumvent the general objections against head movement (including cliticization) raised by Chomsky (2001). The result makes this version of the movement theory of clitics compatible

the verb in (7c). The latter is lexicalized only when it provides the feminine plural reading combining with the occurrence of an OCl *l-a*. This distribution is like that of object agreement on the perfect participle – though in this instance object agreement is (successive-cyclically) marked on the finite verb.

In general, independently of the movement vs. base-generation debate on cliticization, Manzini & Savoia (2017) show that clitic categories match closely those proposed for the sentential spine and that despite apparent idiosyncratic variation, the order of the clitic string corresponds to that of the functional spine of the sentence. Thus dative clitics are fairly routinely taken to be exponents of Appl. The *se/si* clitic is obviously a Voice clitic inducing valency alternations (passive, middle, reflexive, impersonal).

Going back to the structure in (15), OClS may also be deemed responsible for triggering Agree with the perfect participle. In the absence of a clitic, transitivity/accusativity can be introduced by an abstract *v* head. It is tempting to assume that in the absence of the rich ϕ -features set of the clitics, agreement with the participle is not triggered. This would explain the fact that only clitics trigger object-participle agreement, and objects in situ do not, in most Romance languages which have participle agreement.⁴ Uninterpretability of the clitic (the Agree probe) can be maintained under the approach of Sportiche (1996), where the corresponding argument position is filled by a *pro* (moved to the Spec of the clitic head).

Let us turn next to the CP phase. The relation of SClS to finite verbs displays complete parallelism with that of OClS to participles. Thus in the masculine, the plural SCl *i* contrasts with the zero inflection on the finite verb. In the feminine plural, more tellingly, the SCl is *la* and the finite verb bears the *-aŋ* inflections. This suggests that in languages with SClS, it is the SCl, rather than T, that inherits the agreement properties of the C phase head in the CP phase. TP is then the complement of the phase head. Following suggestions on labelling by Chomsky (2013)

with our proposal concerning the phasal head status of the clitic, since the clitic is literally part of the label of *v*.

- (i) a. ...*le voit*
 “he sees him/it”
 b. [_v* *le*_[iφ]] [_v* *voit* [_v* iV, uφ]]

4. D’Alessandro & Roberts (2008) note that standard minimalist theory cannot differentiate object clitics from in situ objects with respect to *v*/perfect participle agreement. Under their account, clitics and perfect participles are in the same Spell-Out domain and hence externalize agreement, while direct objects in situ are isolated within the VP complement of the *v* phase, preventing the externalization of agreement with the perfect participle. We acknowledge both the conceptual similarity to the present account of DPM and the fact that as they now stand, they are mutually incompatible. In this chapter, we suggest that perfect participle agreement in Romance is a core syntax matter, not an externalization matter.

we reinstate Agr(S) or rather ϕP , as a position along the sentential spine, whose content are SCLs in languages like Soazza (or North Italian dialects in general).⁵ Thus the structure of *la dɔrməŋ* “they(f.) sleep” in (3b) is as indicated in (16). (16) is in keeping with the conceptual core of standard analyses of SCLs (Rizzi 1986), where SCLs are conceived essentially as clitic inflections. As in (15), furthermore, we assume that the clitic head triggers Agree with the finite verb.

(16) [ϕP *la* [$_{TP}$ *dɔrm-əŋ*] cf. (3b)

Before concluding, we turn briefly to some language-specific properties of Soazza, namely the fact that not only SCL, but also OCl agrees with the finite verb. This yields distinctive forms such as *la tʃamiəŋ* “I call them(f.)” in (6b’), where the verb base *tʃam-* inflects for the 1st person singular *-i* (*tʃami* “I call”) agreeing with the sentential subject and for the feminine plural *-əŋ* agreeing with the OCl *la*. To begin with, we assume that from the vP phase where they are first-merged, OCls end up in functional positions within the CP phase, as in (17).

(17) [$_{AccVoice}$ *la* [$_{TP}$ [[*tʃam-i*]-*əŋ*]]

At the end of § 2, we mentioned that verbal *-əŋ* is to be identified with nominal *-əŋ*. As such, it contains gender and number specifications, but no Person features. We therefore propose that in (17) the number and gender nominal agreement *-əŋ* can be added on top of a Person inflection (actually a Participant inflection), without one interfering with the other. If the Person inflection agrees with the sentential subject, then the nominal *-əŋ* inflection can still agree with OCl, yielding (17).

In short, *-əŋ* on a finite verb can agree with SCL, as in (16), or with an OCl as in (17), in which case it can be stacked on top of a Person inflection. This means

5. An anonymous reviewer asks us to elaborate on the hypothesis that it is SCLs that host the agreement properties in the CP phase. This assumption fits in with their ability to externalize the agreement properties in many North Italian dialects, where face to the reduced person and number inflection of the verbal paradigms, SCLs supply richer and more systematic phi-features distinctions. The role of SCLs in externalizing phi-features is illustrated by the paradigm of the variety of Soazza itself in (i), where we find a sort of complementary distribution between SCLs, as agreement heads (e.g., 2nd person singular), and specialized inflection (1st singular and 2nd plural).

(i) *dɔrm-i, tu dɔrm, əl/la dɔrm, um dɔrm, dor'm-iŋ, i dɔrm/la dɔrm-əŋ*
 “I sleep, you sleep, etc.”

Informally SCLs and inflections, i.e., the Agr and T/I heads, enter the inheritance mechanism from the phase head C, eventually complementing one another, as in many varieties, where specialized SCLs join with specialized inflections (Manzini & Savoia 2005). However, the specific theoretical question of whether it is possible to conclude that SCLs are the true inheritance head of the CP phase is beyond the scope of the present work.

that in an example like (7a) *la la tfaməŋ* “they(f.) call her” or “she calls them(f.)”, either the SCl or the OCl can trigger the -*əŋ* feminine plural agreement on the finite verb, see (18).

(18) [_{φP} *la* [_{AccVoice} *la* [_{TP} *tfam-əŋ*]] cf. (7a)

As highlighted in the presentation of the data, there is a third possibility, namely that in (18) -*əŋ* agrees with both SCl and OCl determining the plurality of both, namely “They(f.) call them(f.)”. Nevins (2011) briefly considers the Soazza patterns and assumes that the interpretation whereby both OCl and SCl are plural depends on a Multiple Agree configuration, where both the SCl and the OCl enter a probe-goal relation with the same -*əŋ* inflection. Nothing prevents us from assuming this approach.⁶

3.2 Alternative accounts

Before abandoning the issues raised by the asymmetric distribution of plural inflection in the Soazza variety, we highlight some differences between our analysis and those proposed in alternative syntactic accounts.⁷ Generally, syntactic treatments trace back the asymmetry between D and N to mechanisms mediated by the status of number as a functional head Num of DP, projecting NumP in phrasal syntax – and by different types of agreement rules.

In the treatment of Ladin, Rasom (2008), following Giusti (2008), assumes that there is a Concord rule (roughly feature percolation) taking place between N and the functional nodes on the nominal spine. Some languages, like Ladin, are characterized by a restriction whereby Concord takes place only downward (Lazy Concord). Therefore once N moves to the syntactic Num projection, it can percolate its plural feature downwards, hence to its right, but not upwards, hence not to its left (e.g., to Ds).

Stark & Pomino (2009) take up the asymmetries inside the DP in Ladin varieties, as well as in BP varieties, and propose an analysis that, all things considered, is not very different from that of Rasom. For Stark & Pomino gender and

6. Nevins (2011: 8, 9) assumes that the ability of number to extend ambiguously to object or subject descends from the underspecified status of singular, whereby “unmarked values of number, e.g., [-singular], are never syntactically active and never referred to in the syntax”. By contrast, “person features are always fully specified on syntactic arguments”, thus excluding generalization processes.

7. For some discussion of DM treatments, including Costa & Figueredo (2002); Bonet (2018), see § 2 and Savoia et al. (2018, 2019), Manzini et al. (2018).

number agreement depend on two functional head projections of N, namely *n* and Num, forming the hierarchy [NumP [*n*P ... [NP]]]]. In languages like Brazilian Portuguese, where N never displays plural, the *n* head is defective with respect to number. By contrast, number features are present on *n* in languages like Spanish or Ladin, which display plural on N. At the next step of the derivation, Num works as a probe for gender features taking N as its goal: “the probe, once it has its unvalued feature instantiated by its goal gives away its number feature” (p. 127). Once this happens, plural features percolate to *n* in Spanish or Ladin, but not in BP. The result is that in BP N does not realize number, though in Spanish and in Ladin it does. Ladin then ought to be just like Spanish. However, unlike Spanish, feature percolation on pronominal adjectives is avoided because of a constraint whereby “adjectives demand a plural exponent only in phrase final position” (p. 127). The result is circular, given that this restriction is just what should be explained. Furthermore, D is not discussed. Agreement of D in number takes place in BP (p. 130) and evidently in Spanish. Since the only parameter explicitly stated by Stark & Pomino bundles together Ladin and Spanish, an independent parameter seems to be needed to capture lack of number agreement on D in Ladin.

In short, the treatments just reviewed raise descriptive and theoretical questions. They require some ad hoc stipulations, like the restriction of plural to phrase final adjectives just mentioned. Furthermore, the recourse to the interaction of percolation/Concord with Agree (as opposed to just Agree) to explain the distribution of inflectional properties is an enrichment of the model.

An interpretation of split agreement in number within the DP that could in principle be relevant for our analysis is proposed in Landau (2016), where the split agreement in number is illustrated with the Hebrew noun *be'alim* “owner(s)”, morphologically plural, but admitting both plural and singular agreement with the adjective and the verb. Landau adopts the distinction between Index ϕ -features (including person), endowed with referential import, and Concord ϕ -features (including at most gender and number) of a pure syntactic nature. He further divides DP into three agreement zones. Adjectives introduced in the lowest zone probe for N hence for Concord features. Another possibility is that adjectives are introduced in the intermediate zone above Num, giving rise to a different type of agreement including Index features.⁸ The highest one is the D zone: “D is the exclusive contact point between external probes (like *v* and T) and any nominal ϕ -feature” (Landau

8. In (i)–(ii) we reproduce the schemas provided by Landau (2016: 1000) for Concord and Index agreement of the adjective.

- (i) [_{DP} D ... [_{NumP} Num_[INDEX] ... [Adj ... [_{NP} N_[CONCORD]]]]] CONCORD agreement
 (ii) [_{DP} D [Adj [_{NumP} Num_[INDEX] [_{NP} N_[CONCORD]]]]] INDEX agreement

2016: 996). The external probes, *v* and *T*, are specified for Index feature including person, and lack Concord features, and consequently they only target Index features of *D*.

If we compare Landau (2016) to present assumptions, we see a clear difference in the adoption of both Concord and Index features (as opposed to a single ϕ -features set). Again, the richness of the theory is an issue, since Index features encode a notion of abstract referential index, separate from lexical ϕ -features, whose abandonment was a proposal of Chomsky's (1995) minimalist program.⁹ From an empirical point of view, in the phenomena of Soazza examined in this work, *N* clearly lexicalizes the same agreement properties as the verb, to the exclusion of *D*. Therefore there is no clear sense in which *N* would be universally specified for Concord features as opposed to the Index features of *D*. On the other hand, if the agreement involving *D* has semantic import, this could confirm our analysis that in the dialect of Soazza *l-a* is a bona fide specialized form for plural. In other words, it cannot be treated as a default, since if that were the case, the transfer of featural content to the sentential domain would be blocked.

4. Conclusions

Many non-standard Romance varieties are characterized by the phenomenon of Differential Plural Marking, illustrated in § 1 with data from North Lombard varieties of the Mesolcina Valley. Under DPM, phasal domains have a different realization of plural morphology on the head of the phase and on the complement of the phase. We provided an account of this split in § 3, based on the idea that the complement of the phase and its head are externalized separately. We have drawn consequences concerning clitics as phasal heads as well as object agreement with participles and with finite verbs.

9. A full discussion of this point is beyond the scope of the present work. Relevant data are provided by the so-called neuter of Italo-Albanian dialects, involving *Ns* with mass denotation, where the plural morphology of the noun, of demonstratives, linkers and adjectives is matched with the singular morphology of the verb. The account proposed by Manzini & Savoia (2018b); Baldi & Savoia (2018) is that DP-internal categories agree with respect to [aggregate] features (see § 2), encompassing both mass and count plural. On the contrary, the verb is sensitive only to count plural and hence agrees in the singular. High and low plurals are specifically addressed by Manzini (2020).

References

- Acquaviva, Paolo. 2008. *Lexical plurals*. Oxford: Oxford University Press.
- Alexiadou, Artemis. 2011. Plural mass nouns and the morpho-syntax of number. In Mary B. Washburn, Katherine McKinney-Bock, Erika Varis, Ann Sawyer, & Barbara Tomaszewicz (eds.), *Proceedings of WCCFL 28*, 33–41. Somerville, MA: Cascadia Proceedings Project.
- Baldi, Benedetta & Ludovico Franco. 2018. Gender, number and inflectional class in some North-Italian dialects. The plural inflection *-i* and the interpretation of N morphology. In Mirko Grimaldi, Rosangela Lai, Ludovico Franco & Benedetta Baldi (eds.), *Structuring variation in Romance Linguistics and beyond*, 31–54. Amsterdam: John Benjamins.
<https://doi.org/10.1075/la.252.02bal>
- Baldi, Benedetta & Leonardo M. Savoia. 2018. Morphosyntactic reorganization phenomena in Arbëresh dialects: The neuter. *Quaderni di Linguistica e Studi Orientali* 4. 109–130.
- Bonet, Eulàlia. 2018. Missing inflectional features and missing exponents in DP-internal agreement asymmetries. *Glossa* 3(1). 79. 1–19. <https://doi.org/10.5334/gjgl.579>
- Carstens, Vicky. 2000. Concord in Minimalist Theory. *Linguistic Inquiry* 31. 319–355.
<https://doi.org/10.1162/002438900554370>
- Chierchia, Gennaro. 2010. Mass nouns, vagueness and semantic variation. *Synthese* 174. 99–149.
<https://doi.org/10.1007/s11229-009-9686-6>
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–54. Cambridge, MA: The MIT Press.
- Chomsky, Noam. 2005. Three factors in language design. *Linguistic Inquiry* 36. 1–22.
<https://doi.org/10.1162/0024389052993655>
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos Otero & Maria-Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory. Essays in honour of Jean-Roger Vergnaud*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9780262062787.003.0007>
- Chomsky, Noam. 2013. Problems of projection. *Lingua* 130. 33–49.
<https://doi.org/10.1016/j.lingua.2012.12.003>
- Costa, João & M. Cristina Figueiredo Silva. 2002. Notes on nominal and verbal agreement in Portuguese. *Rivista di Grammatica Generativa* 27. 17–29.
- D'Alessandro, Roberta & Ian Roberts. 2008. Movement and agreement in Italian past participles and defective phases. *Linguistic Inquiry* 39(3). 477–491.
<https://doi.org/10.1162/ling.2008.39.3.477>
- Déchine, Rose-Marie, Raphaël Girard, Calisto Mudzingwa & Martina Wiltschko. 2014. The internal syntax of Shona class prefixes. *Language Sciences* 43. 18–46.
<https://doi.org/10.1016/j.langsci.2013.10.008>
- Giusti, Giuliana. 2008. Agreement and concord in nominal expressions. In Cécile de Cat & Katherine Demuth (eds.), *The Bantu–Romance connection: A comparative investigation of verbal agreement, DPs, and information structure*, 201–237. Amsterdam: John Benjamins.
<https://doi.org/10.1075/la.131.12giu>
- Higginbotham, James. 1985. On semantics. *Linguistic Inquiry* 16. 547–621.
- Kayne, Richard. 1991. Romance clitics, verb movement and PRO. *Linguistic Inquiry* 22. 647–686.
- Kramer, Ruth. 2016. A split analysis of plurality: Number in Amharic. *Linguistic Inquiry* 47(3). 527–559. https://doi.org/10.1162/LING_a_00220

- Landau, Idan. 2016. DP-internal semantic agreement: A configurational approach. *Natural Language and Linguistic Theory* 34. 975–1020. <https://doi.org/10.1007/s11049-015-9319-3>
- Manzini, M. Rita. 2020. The morphosyntactic structure of number in Italian and Albanian. High and low plurals. *Catalan Journal of Linguistics* 19. 127–157. <https://doi.org/10.5565/rev/catjl.319>
- Manzini, M. Rita, Benedetta Baldi & Leonardo M. Savoia. 2018. Plural morphology in DPs in -s and -i systems: A Romance comparative perspective. In Annamaria Chillà & Alessandro De Angelis (eds.), *Capitoli di morfossintassi delle varietà romanze d'Italia: teoria e dati empirici*, 307–330. Palermo: Supplementi del Bollettino del Centro Studi Filologici e Linguistici Siciliani.
- Manzini, M. Rita & Leonardo M. Savoia. 2005. *I dialetti italiani e romanci. Morfosintassi generativa*. Alessandria: Edizioni dell'Orso.
- Manzini, M. Rita & Leonardo M. Savoia. 2011. Mesoclis in the imperative: Phonology, morphology or syntax? *Lingua* 121. 1101–1120. <https://doi.org/10.1016/j.lingua.2011.02.002>
- Manzini, M. Rita & Leonardo M. Savoia. 2017. Enclisis/Proclisis alternations in Romance: Allomorphies and (re)ordering. *Transactions of the Philological Society* 115. 98–136. <https://doi.org/10.1111/1467-968X.12093>
- Manzini, M. Rita & Leonardo M. Savoia. 2018a. N inflections and their interpretation: Neuter -o and plural -a in Italian varieties. In Roberto Petrosino, Pietro Cerrone & Harry van der Hulst (eds.), *From sounds to structures: Beyond the veil of Maya*, 361–394. Berlin: Mouton De Gruyter.
- Manzini, M. Rita & Leonardo M. Savoia. 2018b. *The morphosyntax of Albanian and Aromanian varieties*. Berlin: Mouton De Gruyter. <https://doi.org/10.1515/9781501505140>
- Manzini, M. Rita & Leonardo M. Savoia. 2019. N morphology and its interpretation: Romance feminine singular/plural -a. In Silvio Cruschina, Adam Ledgeway & Eva-Maria Remberger (eds.), *Italian dialectology at the interfaces*, 257–293. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.251.12man>
- Nevins, Andrew. 2011. Multiple agree with clitics: Person complementarity vs. omnivorous number. *Natural Language & Linguistic Theory* 29(4). 939–971. <https://doi.org/10.1007/s11049-011-9150-4>
- Rasom, Sabrina. 2008. Lazy concord in the central Ladin feminine plural DP. Ph.D. Dissertation: Università di Padova.
- Rizzi, Luigi. 1986. On the status of subject clitics in Romance. In Osvaldo Jaeggli & Carmen Silva-Corvalan (eds.), *Studies in Romance linguistics*, 391–419. Dordrecht: Foris. <https://doi.org/10.1515/9783110878516-025>
- Roberts, Ian. 2010. *Agreement and head movement: Clitics, incorporation, and defective goals*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9780262014304.001.0001>
- Rohlf, Gerhard. 1968 [1949]. *Grammatica storica della lingua italiana e dei suoi dialetti. Morfologia*. Torino: Einaudi.
- Salvioni, Carlo. 1902. Del plurale femminile di 1. declinazione esposto per -a ed -an in qualche varietà alpina di Lombardia: nota. *Rendiconti del Reale Istituto lombardo di scienze e lettere* 2(35). 905–919.
- Savoia, Leonardo M., Benedetta Baldi & M. Rita Manzini. 2018. Sigmatic plurals in Romance varieties spoken in Italy and their interaction with -i plurals. *LINGBAW* 4. 141–160. <https://doi.org/10.31743/lingbaw.5671>

- Savoia, Leonardo M., Benedetta Baldi & M. Rita Manzini. 2019. Asymmetries in plural agreement in DPs. In Joseph Emonds, Markéta Janebová, & Ludmila Veselovská (eds.), *Language use and linguistic structure Proceedings of the Olomouc Linguistics Colloquium 2018*, 204–224. Olomouc: Palacký University Olomouc.
- Sportiche, Dominique. 1996. Clitic constructions. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 213–276. Dordrecht: Kluwer.
https://doi.org/10.1007/978-94-015-8617-7_9
- Stark, Elisabeth & Natascha Pomino. 2009. Adnominal adjectives in romance. Where morphology seemingly meets semantics. In M. Teresa Espinal, Manuel Leonetti & Louise McNally (eds.), *Proceedings of the IV Nereus International Workshop: Definiteness and DP structure in Romance languages*, 113–135. Konstanz, Germany.

Brazilian and European Portuguese and Holmberg's 2005 typology of null subject languages

Ana Maria Martins and Jairo Nunes
Universidade de Lisboa / Universidade de São Paulo

This paper rethinks Holmberg's (2005) characterization of partial vs. consistent null subject languages (NSL) based on data from Brazilian and European Portuguese, the former a partial, the latter a consistent NSL. The paper proposes that rather than overt morphological distinctions, what is relevant for null subject licensing is the underlying feature specification of the verbal inflection, after agreement between T and a pronominal subject values the relevant person/number/gender/Case feature. Hence, only close inspection of the pronominal and agreement systems of individual NSLs permits an adequate characterization of them, for the same language may behave as a 'partial', 'consistent', or 'radical' NSL depending on the morphological feature specification of its nominative pronouns and T heads.

Keywords: null subjects, consistent/partial *pro*-drop languages, ellipsis, verbal agreement morphology, pronominal system, feature valuation, feature prominence, person, number, gender, case

1. Consistent, partial, and radical *pro*-drop languages: Questions and aim of the paper

Brazilian Portuguese (BP) and European Portuguese (EP) have been respectively analyzed as canonical examples of partial and consistent null subject languages (NSLs) within Holmberg's (2005) typological tripartition of NSLs into 'consistent', 'partial', and 'radical' (or 'discourse') *pro*-drop languages.¹ According to Holmberg

1. For relevant discussion and further elaboration on this tripartite division, see e.g., Holmberg, Nayudu & Sheehan 2009; Biberauer et al. 2009; Holmberg 2009a; b, 2016; Sheehan 2016; Barbosa 2019.

(2005, 2009a; b, 2016), pro-drop is dependent on agreement (or control) in consistent and partial NSLs, but not in radical NSLs (which typically have no verbal agreement). Consistent NSLs are taken to involve a rich system of subject-verb agreement and unrestricted pro-drop (affecting all persons, all clause types, main and embedded clauses, initial and non-initial positions, and all registers), whereas partial NSLs are taken to involve a deficient subject-verb agreement system and restricted pro-drop (affecting only some persons/verb forms or being sensitive to differences of clause type, main/embedded configuration, or register). Holmberg (2009a; b) and Roberts (2010) propose that these differences follow from a difference in the featural content of T, which is taken to bear a D(efiniteness)-feature in consistent but not in partial NSLs. The presence of the D-feature on T in consistent NSLs triggers the interpretation of a null subject as a definite pronoun, thus excluding the interpretation of a third person singular null subject as an inclusive generic subject (corresponding to *one* in English). In turn, since T does not bear a D-feature in partial pro-drop languages, their null subjects can only receive a definite reading under control by a local antecedent and a third person singular null subject may have the inclusive generic reading.

Contrary to the predictions of Holmberg's (2009a; b) theory, EP allows sentences like (1) below, with a third person singular null subject with an inclusive generic reading, but displays all other characteristic properties of consistent NSLs.

- (1) *Meu querido, isto aqui é assim: Ø deitou, Ø pagou.*
 my dear this here is so laid.down paid
 "My dear, this is how it works: once one has laid down, one will have to pay for the session."
 (said by a physiotherapist)

The availability of (1) in EP can be accounted for if we extend to EP Carvalho's (2019) analysis of indefinite third person singular null subjects in BP. Carvalho convincingly argues that these null subjects actually involve two different types: a ϕ -defective one, which receives an existential interpretation (see (2) below), and a ϕ -complete one, which receives a generic interpretation (see (3)).² The ϕ -defective existential null indefinite is licensed by a finite T-head bearing only a number feature (a possibility available in BP but not in EP; see e.g., Ferreira 2000, 2009; Nunes 2008), whereas the pronoun in generic null impersonals is licensed by a regular finite T, with both number and person features. Therefore, Carvalho's proposal correctly predicts that EP should exclude (2) but could allow (3), if Holmberg's D-feature is discarded.

2. Sentence (2) is ungrammatical in EP only under the indefinite reading indicated by the translation; under a definite reading, where the null subject takes the matrix subject or some other salient element in the discourse for its antecedent, the sentence is perfectly fine.

- (2) *Ele disse que Ø vende sopa em qualquer restaurante.* (EP: *; BP: OK)
 he said that sells soup in any restaurant
 "He said that one sells soup in any restaurant."
- (3) *Antigamente era assim: Ø falava mais de uma língua, Ø 'tava*
 in.the.old.days was so spoke more of one language was
empregado (EP/BP: OK)
 employed
 "In the old days, it was like this: anyone who spoke more than one language
 would have a job."

In this paper, we deal with other additional empirical facts revealed by EP and BP that do not seem to be fully captured by Holmberg's (2009a; b) and similar theories of pro-drop (see Sheehan 2016 for an overview and references).³ We specifically consider the following four questions, which pose challenges to Holmberg's characterization of consistent versus partial NSLs: (i) Why do null subjects in BP display different degrees of acceptability depending on the pronoun that is 'dropped' (see (4) below)?; (ii) Why is the null subject corresponding to *a gente* "we" unacceptable in both BP and EP (see (4e))? (Note that there is no simple correlation between the overt realization of verbal agreement and the pattern of (un)grammaticality; cf. (4a) and (4e)); (iii) Why are null subjects without a local 'controller' allowed in participial clauses in the absence of 'rich agreement' (see the sentences in (5),

3. Since our goal here is to investigate the role played by agreement in pro-drop licensing, we set aside syntactic configurations of 'controlled' subjects (see (i)) and discourse/syntactic contexts allowing topic-drop (see (iia), displaying topic-drop, versus (iib), excluding it).

- (i) a. *[A gente]_i acha que [Ø]_i deve participar mais*
 we think.Ø that should.Ø participate more
 "We think that we should be more participative." (Ø = *a gente* → EP/BP: OK)
- b. *[o João]_i disse que [o Pedro]_k acha que Ø_{k/i} vai ser promovido*
 the João said that the Pedro thinks that goes be promoted
 "João_i said that [Pedro]_k thinks that he_{k/i} is going to be promoted"
 (Ø = *o Pedro* → EP/BP: OK; Ø = *o João* → EP: OK; BP: *)

(ii) Context

Onde está a Maria?

where is the Maria
 'Where's Maria?'

- a. *[Ø acabou de sair]* (Ø = *ela* → EP/BP: OK)
 finished.3SG of leave
 "She's just left."
- b. *O que Ø fez desta vez?* (Ø = *ela* → EP: OK; BP: *)
 the what did.3SG of-this time
 "What did she do this time?"

which are fine in both BP and EP)?; and (iv) Why are noncontrolled null subjects allowed in uninflected gerund clauses in BP and EP (see (6)), which in this respect apparently behave like radical pro-drop languages?

- (4) a. [O que Ø quer fazer]?
 the what want.Ø do
 “What {do you(SG) / does he} want to do?”
 (Ø = *você* (you(SG)) / *ele* (he) → EP: OK; BP: *)
- b. [O que Ø tenho a ver com isso]?
 the what have.1SG to see with this
 “What do I care? / What do I have to do with it?”
 (Ø = *eu* (I) → EP: OK; BP: ??)
- c. [Quando Ø vão viajar]?
 when go.3PL travel
 “When will you(PL) be travelling?”
 Ø = *vocês* (you(PL)) / *eles* (they) → EP: OK; BP: ??)
- d. [Quem Ø devemos contratar]?
 who should.1PL hire
 “Who should we hire?” (Ø = *nós* (we) → EP/BP: OK)
- e. [Quando Ø deve viajar]?
 when should.Ø travel
 “When are we supposed to travel?” (Ø = *a gente* (we) → EP/BP: *)
- (5) *Ninguém esperava muito d[o João]_i. Mas [depois de Ø_i nomeado para o cargo], a empresa melhorou consideravelmente.*
 nobody expected much of.the João but after of
 appointed.MASC.SG to the position the company improved
 considerably
 “Nobody expected much of João. But after his being appointed to the position, the company considerably improved.” (EP/BP: OK)
- (6) a. [Ø Agindo honestamente], todos vão votar em você.
 acting honestly all go vote in you
 “If you do things honestly, everyone is going to vote for you.”
 (Ø = *você* → EP/BP: OK)
- b. [Ø Usando roupas adequadas], o frio é suportável.
 using clothes adequate the cold is bearable
 “If {you wear / one wears} proper clothes, the cold is bearable.”
 (Ø = *você* → EP/BP: OK)

The paper is organized as follows. In § 2, we put forward a formal proposal to address the questions raised above and show how it works in finite domains (see

questions (i) and (ii) above). The proposal is then extended to participial clauses in § 3 (see question (iii)) and gerund clauses in § 4 (see question (iv)). Section 5 concludes the paper.

2. A fine-grained investigation on verbal agreement and null subject licensing in BP and EP

Perlmutter's (1971) deletion approach to null subjects was revived within Minimalism by different authors (see Sheehan 2006, 2016; Saab 2008, 2016; Roberts 2010; Duguine 2013, among others). Subscribing to this general approach, we propose that pronominal subject deletion is governed by the condition in (7). Specifically, pronominal subject ellipsis is available in languages with a positive setting for subcondition (7i) and is licensed under subcondition (7ii), where the relevant scale of prominence is given in (8), with features ordered from more to less salient.⁴

(7) Prominent Feature Valuation Condition on pronominal subject ellipsis

T can license ellipsis of a pronominal subject *pro* if:

- i. T is associated with an E(llipsis)-feature (e.g., Merchant 2001); and
- ii. agreement between T and *pro* involves valuation of the most prominent feature of T.

(8) Feature prominence: *person* > *number* > *gender* > *Case*

4. The E-feature in (7i) is simply a descriptive statement saying that some heads (in this case T) are able to license ellipsis and some heads are not. At this point, we put aside the issue of which heads can license ellipsis in a given language and focus on the condition in (7ii). In Merchant's (2001) system, the E-feature triggers ellipsis of the complement of the head that bears E. In the case at hand, we assume that a T head bearing an E-feature may license ellipsis of the DP it agrees with. That being so, one reviewer asks whether we are predicting correlations between these two types of ellipsis. For instance, could a head bearing an E-feature license both ellipsis of its complement and ellipsis of a pronoun it agrees with? We assume that these two types of ellipsis can be subsumed under a single structural relation but are associated with different licensing conditions. That is, both types may instantiate a probe-goal relation, with the head bearing E probing its (c-command) domain: in the case of complement ellipsis, the head bearing E symmetrically c-commands the elided complement and in the case of pronominal subjects, ellipsis may be licensed while the subject is within vP and the local T_E asymmetrically c-commands it (Alternatively, if C is the host of clausal ϕ -features as in Chomsky 2008, pronominal subjects may be elided in Spec,TP if the local C bears E). The licensing condition is selection in the first case and the prominence requirement in (7ii) in the second case. Thus, pending further unification between selection and agreement (see e.g., Wurmbrand 2014), we do not anticipate any direct correlation between these two types of licensing and accordingly, between these two types of ellipsis.

Assuming that both BP and EP have a positive setting for (7i), their differences arise from the different feature composition of their pronouns (see Nunes 2019) and the corresponding agreement inflection they are associated with. This is shown in Table 1 (see (4) above for illustration).⁵

Table 1. Feature composition of pronouns in BP and EP and null subject licensing infinite clauses

Nominative pronouns	Morphological specifications		Valuation of T after agreement		Null subject in finite clauses	
	EP	BP	EP	BP	EP	BP
<i>nós</i> “we”	[P/N:1.PL]	[P/N:1]	[P/N:1.PL]	[P/N:1]	OK	OK
<i>eu</i> “I”	[P/N:1.SG]	[P/N:SG]	[P/N:1.SG]	[P/N:SG]	OK	??
<i>vocês</i> “you.PL”	[P:2; N:PL]	[P; N:PL]	[P:2; N:PL]	[P:u; N:PL]	OK	??
<i>eles</i> “they.MASC”	[P:3; G:MASC; N:PL]	[P; G:MASC; N:PL]	[P:3; N:PL]	[P:u; N:PL]	OK	??
<i>elas</i> “they.FEM”	[P:3; G:FEM; N:PL]	[P; G:FEM; N:PL]	[P:3; N:PL]	[P:u; N:PL]	OK	??
<i>tu</i> “you.SG”	[P/N:2.SG]	%[P; N]	[P/N:2.SG]	([P:u; N:u])	OK	*
<i>você</i> “you.SG”	[P:2; N:SG]	[P; N]	[P:2; N:SG]	[P:u; N:u]	OK	*
<i>ele</i> “he”	[P:3; G:MASC; N:SG]	[P; G:MASC; N]	[P:3; N:SG]	[P:u; N:u]	OK	*
<i>ela</i> “she”	[P:3; G:FEM; N:SG]	[P; G:FEM; N]	[P:3; N:SG]	[P:u; N:u]	OK	*
<i>a gente</i> “we”	[P/N]	[P/N]	[P:u; N:u]	[P:u; N:u]	*	*

As Table 1 shows, the acceptability of null subjects in each language is determined by the hierarchy in (8). Therefore, the best result is when T is valued in person and the second best result is when T is valued in number. All the remaining cases lead to unacceptable results in finite clauses, as neither person nor number is valued. In EP all the pronouns but *a gente* “we” are specified for person; hence, only *a*

5. The pronoun *tu* “you.2sg” is dialectally restricted in BP (annotated by “%” in Tables 1 and 2). In the dialects where it is used productively, it typically triggers the same agreement pattern as *você*. Accordingly, *tu* and *você* in BP were assigned the same morphological specifications in Table 1. Formal registers of these dialects that maintain an EP-style agreement for *tu* can be accounted for in the present system if *tu* is morphologically specified as [P/N:2], coupled with the following correspondence rule for the realization of the agreement inflection it triggers: [P/N:2] > -ste / IND.PERFEC.PAST and -s elsewhere (see Table 3 below). For purposes of exposition, we will put these formal registers aside.

gente (which is the diachronic outcome of the grammaticalization of the nominal expression “the people”) disallows a corresponding null subject.

The diachronic change that BP underwent from a consistent to a partial NSL has been related to changes in the pronominal system leading to the weakening of subject-verb agreement morphology (e.g., Duarte 1993, 2000; Galves 1993). However, agreement morphology by itself cannot account for grammaticality contrasts such as the ones in (9)–(11) below, which show no clear correlation with surface verbal morphology.

- (9) EP: *a gente* “we” vs. *você/ele* “you.SG/he”]
Eles pensam que [Ø não quer reclamar]
 they think that not want.Ø complain
 “They think that {we do / you do / he does} not want to complain.”
 (Ø = *a gente* → *; Ø = *você/ele* → OK)
- (10) [*você/ele* “you.SG/he”: EP vs. BP]
Eles pensam que [Ø não quer reclamar]
 they think that not want.Ø complain
 “They think that {you do / he does} not want to complain.”
 (Ø = *você/ele* → EP: OK; BP: *)
- (11) [BP/EP: *eu* “I” vs. *nós* “we”]
 a. *O professor disse que [Ø escrevo bem]*
 the teacher said that write.1(SG) well
 “The teacher said that I write well.” (Ø = *eu* → EP: OK; BP: ??)
 b. *O professor disse que [Ø escrevemos bem]*
 the teacher said that write.(1)PL well
 “The teacher said that we write well.” (Ø = *nós* → EP/BP: OK)

In fact, BP and EP do not significantly differ with respect to the overt marking of verbal agreement inflection (see the last two columns of Table 2 below). Therefore, their different behavior relative to the licensing of null subjects indicates that the same phonological exponent may be the realization of different morphological specifications in each language, as shown in Tables 2 and 3.

Observe that the last two columns of Table 2 show an interesting difference between the verbal inflection associated with the indicative present tense and the one associated with the indicative imperfective past: the former overtly marks agreement with *eu*, but the latter does not. We take the imperfective past to be subject to a process of morphological impoverishment (in the sense of Bonet 1991), which deletes the number value of inflection associated with *eu*, thus activating the elsewhere correspondence rules in Table 3. We therefore predict that morphological impoverishment should have an impact on null subject licensing in BP, but not in EP. The reason is that impoverishment leaves a valued person feature in T in EP, but

Table 2. Morphological specifications for person and number and verbal agreement in Portuguese

Nominative pronouns	Morphological specification for person and number				Surface form of <i>dançar</i> “dance”	
	EP		BP		INDIC. PRES.	INDIC. IMPERF. PAST
	Pronoun specification	Agreement inflection	Pronoun specification	Agreement inflection		
<i>nós</i> “we”	[P/N:1.PL]	[P/N:1.PL]	[P/N:1]	[P/N:1]	<i>dançamos</i>	<i>dançávamos</i>
EP <i>tu</i> “you.SG”	[P/N:2.SG]	[P/N:2.SG]			<i>danças</i>	<i>dançavas</i>
<i>eu</i> “I”	[P/N:1.SG]	[P/N:1.SG]	[P/N:SG]	[P/N:SG]	<i>danço</i>	
%BP <i>tu</i> “you.SG”						
<i>ocê</i> “you.SG”	[P:2; N:SG]	[P:2; N:SG]	[P; N]	[P:u; N:u]	<i>dança</i>	<i>dançava</i>
<i>ele/ela</i> “he/she”	[P:3; N:SG]	[P:3; N:SG]				
<i>a gente</i> “we”	[P/N]	[P:u; N:u]	[P/N]			
<i>vocês</i> “you.PL”	[P:2; N:PL]	[P:2; N:PL]				
<i>eles/elas</i> “they. MASC/FEM”	[P:3; N:PL]	[P:3; N:PL]	[P; N:PL]	[P:u; N:PL]	<i>dançam</i>	<i>dançam</i>

Table 3. Correspondence rules for the morphological realization of verbal agreement inflection

EP	BP
a. [P/N:1.PL] ↔ {-mos}; [N:PL] ↔ {-m}	a. [P:1] ↔ {-mos}; [N:PL] ↔ {-m}
b. [P/N:1.SG] → {-o} / INDIC.PRES __ → {-i} elsewhere.	b. [N:SG] → {-o} / INDIC.PRES __ → {-i} / INDIC.PERF.PAST __
c. [P/N:2.SG] → {-ste} / INDIC.PERF.PAST __ → {-s} elsewhere.	c. ∅ elsewhere.
d. ∅ elsewhere.	

no valued feature in BP. Hence, an impoverished T should still license a null subject in consonance with the Prominent Feature Valuation Condition (see (7ii)) in EP, but not in BP (not even marginally). The data in (12) shows that the prediction is indeed borne out.

- (12) a. *Todos acham [que ∅ trabalho na fábrica]*
everyone think that work.INDIC.PRES in.the factory
“Everyone thinks that I work at the factory.” (∅ = *eu* → EP: OK; BP: ??)
- b. *Todos achavam [que ∅ trabalhava na fábrica]*
everyone thought that work.INDIC.IMPERF in.the factory
“Everyone thought that I worked at the factory.” (∅ = *eu* → EP: OK; BP: *)

This comparative analysis of EP and BP verbal agreement systems thus allows us to give clear, formal content to the general observation that subject-verb agreement is deficient in BP. It also captures the correlation between the diachronic change towards a partial NSL and the weakening of the verbal agreement morphology in BP (e.g., Duarte 1993, 2000; Galves 1993). However, this correlation is taken to hold at an abstract level, being sensitive to the morphological specifications of the relevant verbal agreement inflections, rather than their phonological exponents.⁶

Let us now reconsider the behavior of *a gente*. Recall that the null counterpart of *a gente* cannot be licensed in either BP or EP (see (4e) and (9)). In a sense, it is as if EP, in this particular case, ceases to be a consistent NSL. Our analysis derives this fact in a straightforward fashion. Given that *a gente* is morphologically underspecified for person and number in both BP and EP (see Table 1), it cannot value the person and number features of T and therefore, its deletion violates the Prominent Feature Valuation Condition in (7ii). This proposal also allows us to account for the diverse agreement patterns involving *a gente* found in dialectal/colloquial EP. Once *a gente* has no intrinsic morphological value for person or number, different dialects/register may then resort to (complete or partial) semantic agreement instead of default agreement, as illustrated in (13) and (14).

6. From this perspective, the distance between standard BP and the varieties/register with fewer agreement distinctions, as exemplified in (i), is in fact much shorter than usually thought. The paradigm in (i), for instance, can be derived in a simple manner if morphological impoverishment deletes the values 1 and PL in Table 2.

- (i) [BP dialects/colloquial registers]
- a. *No final eu sempre acabo concordando.* ([N:SG] → {-o} / INDIC.PRES _)
in.the end I always **finish.SG** agreeing
“In the end I always end up agreeing.”
 - b. *No final nós sempre acaba concordando.* (deletion of 1 → ∅ elsewhere)
in.the end we always **finish.∅** agreeing
“In the end we always end up agreeing.”
 - c. *Vocês não sabe o que aconteceu!* (deletion of PL → ∅ elsewhere)
you.PL not **know.∅** the what happened
“You(PL) don’t know what happened.”
 - d. *Eles nunca pensa em ninguém.* (deletion of PL → ∅ elsewhere)
they never **think.∅** in nobody
“They never think about people.”

Our analysis predicts that in the BP dialect exemplified above only the sentence in (ia) will marginally allow deletion of the semantically first person singular pronoun, due to its preserved morphological specification for number.

- (13) a. *A gente chegou esfomeado/esfomeada.* (EP/BP)
 we arrived starving.MASC/FEM.N:U
- b. *A gente chegou esfomeados/esfomeadas.* (EP)
 we arrived starving.MASC/FEM.PL
 “We arrived starving.”
- (14) a. *Agora estes rapazes novos não sabem o que a gente passámos.*
 now these boys new not know the what we passed.IPL
 “These youngsters have no idea of all we’ve been through.” (%EP)
 (CORDIAL-SIN, Câmara de Lobo e Caniçal, Funchal)
- b. *Que a gente tosquiavam as ovelhas duas vezes no ano.* (%EP)
 that we sheared.PL the sheep two times in.the year
 “Because we used to shear the sheep twice a year.”
 (CORDIAL-SIN, Ponta Graça, Ponta Delgada)

Last but not least, the present proposal eliminates undesirable mismatches between semantic and morphological specifications in the pronominal systems of both BP and EP. The pronouns *ocê* and *a gente*, for example, are commonly described as triggering third person singular agreement, despite the fact that from a semantic point of view, the former is a second person singular pronoun and the latter a first person plural pronoun. As shown in Table 4, there is no mismatch between morphological and semantic specifications. Apparent mismatches are actually due to the interaction of morphological underspecification (see Table 1) with default agreement (the elsewhere case for the phonological realization of inflectional agreement morphemes in Table 3).⁷

7. A reviewer asks how our proposal that *a gente* does not have intrinsic morphological values for person and number is compatible with the “agreement facts on locality discussed by Menuzzi (2000) and Costa & Pereira (2013), since these authors show that different agreement patterns interact with locality conditions”. Menuzzi (2000), mainly dealing with binding, and Costa & Pereira (2005), considering predicative constructions, propose that the different agreement patterns found with the pronoun *a gente* result from agreement being established with its morphological or its semantic features. Typically, the former pattern emerges locally, while the latter emerges non-locally. Costa & Pereira (2013), on the other hand, treat the instances where *a gente* (locally) induces plural verbal agreement as structures with *a gente* generated in apposition to an empty pronoun with first person plural features. In the current paper we don’t have anything to say with respect to (non-local) binding agreement. As for verbal agreement, the dialectal variation between default morphological agreement and agreement induced by semantic features is interpreted as an effect of the ϕ -feature underspecification of *a gente*, as discussed above.

Table 4. Feature composition of personal pronouns in EP and BP:
Eliminating feature mismatch

Nominative pronouns	Person and number		
	Semantic specification	Morphological specification EP	Morphological specification BP
<i>eu</i>	[P:1; N:SG]	[P/N:1.SG]	[P/N:SG]
<i>você</i>	[P:2; N:SG]	[P:2; N:SG]	[P; N]
<i>ele, ela</i>	[P:3; N:SG]	[P:3; N:SG]	[P; N]
<i>nós</i>	[P:1; N:PL]	[P/N:1.PL]	[P/N:1]
<i>a gente</i>	[P:1; N:PL]	[P/N]	[P/N]
<i>vocês</i>	[P:2; N:PL]	[P:2; N:PL]	[P; N:PL]
<i>eles, elas</i>	[P:3; N:PL]	[P:3; N:PL]	[P; N:PL]

3. Null subjects without ‘rich agreement’: Participial clauses

In both EP and BP, participial clauses may display overt nominative subjects, as exemplified in (15) with the third person nominative pronoun *eles* ‘they’.

- (15) a. *Com os cupcakes prontos, cortá-los com um aro. Depois*
with the cupcakes ready cut-them with a hoop after
deles cortados, rechear por dentro e por fora
of.they.MASC.PL cut.PPLE.MASC.PL stuff by inside and by outside
‘‘With the cupcakes ready (baked and cooled), cut them using a hoop. After they are cut, fill and frost them.’’
(<https://issuu.com/boxvox/docs/revista-dio-alterada> (30 May, 2019))

As the data in (16) below show, both varieties of Portuguese also allow noncontrolled third person null subjects in participial clauses. Interestingly, past participles do not display ‘rich agreement’, as they lack person specifications.

- (16) a. *Ninguém esperava muito d[o João]; Mas [depois de Ø,*
nobody expected much of.the João but after of
nomeado para o cargo], a empresa melhorou
appoint.PPLE.MASC.SG to the position the company improved
consideravelmente (EP/BP: OK)
considerably
‘‘Nobody expected much of João. But after his being appointed to the position, the company considerably improved.’’

- b. *Havia uma grande controvérsia sobre [a nova lei]_i. Mas [uma vez Ø_i had a big controversy about the new law but one time votada], ninguém mais prestou atenção ao assunto.*
 vote.PP.LE.FEM.SG nobody more paid attention to.the subject
 “There was a great controversy about the new law. But once it was voted, nobody paid attention to the issue anymore.” (EP/BP: OK)
- c. *Todos só falavam d[os novos funcionários]_i. [Depois de Ø_i all only spoke of.the new employees after of contratados], a produção dobrou.* (EP/BP: OK)
 hire.PP.LE.MASC.PL the production doubled
 “Everybody only talked about the new employees. After they were hired, production doubled.”
- d. *Todos esperavam pel[as novas vagas de trabalho]_i. [Mas depois all waited for.the new vacancies of work but after de Ø_i anunciadas], o diretor cancelou o concurso.*
 of announce.PP.LE.FEM.PL the director canceled the tender
 “Everyone was waiting for the new opening positions. But after they were announced, the director canceled the tender.” (EP: OK; BP: OK)

Once participial Ts lack person, number should count as the most prominent feature in the feature prominence scale in (8) (*person > number > gender > Case*) and its valuation should be taken into account for purposes of null subject licensing, as determined by the Prominent Feature Valuation Condition in (7), repeated here in (17).

(17) **Prominent Feature Valuation Condition on pronominal subject ellipsis**

T can license ellipsis of a pronominal subject *pro* if:

- i. T is associated with an E(llipsis)-feature (e.g., Merchant 2001); and
- ii. agreement between T and *pro* involves valuation of the most prominent feature of T.

BP differs from EP in this regard, though. As shown in (18), number agreement on participial T is not categorically required in (colloquial) BP, while gender agreement is (see, e.g., Simioni 2011). This indicates that due to weakening of number agreement in BP, gender counts as the most prominent feature in the feature prominence scale for BP participials.

- (18) a. *As proposta foram analisada.* (BP: %; EP: *)
 the.FEM.PL proposal.FEM.SG were analyzed.FEM.SG
- b. **As proposta foram analisados* (BP/EP: *)
 the.FEM.PL proposal.FEM.SG were analyzed.MASC.PL
 "The proposals were analyzed."

The fact that the most prominent feature of participials is number in EP but gender in BP does not have a noticeable effect for the licensing of third person null subjects. Once valuation of the most prominent feature of participial T in each language is ensured, as shown in Table 5, a third person null subject is always licensed, as seen in (16).

Table 5. Participial agreement and third person null subjects

3rd person nominative pronouns	EP			BP		
	Morphological specification for gender and number	Participial inflection	Null subject licensing	Morphological specification for gender and number	Participial inflection	Null subject licensing
<i>ele</i>	[G:MASC; N:SG]	[G:MASC; N:SG]	OK	[G:MASC; N]	[G:MASC; N:u]	OK
<i>ela</i>	[G:FEM; N:SG]	[G:FEM; N:SG]	OK	[G:FEM; N]	[G:FEM; N:u]	OK
<i>eles</i>	[G:MASC; N:PL]	[G:MASC; N:PL]	OK	[G:MASC; N:PL]	[G:MASC; N:PL]	OK
<i>elas</i>	[G:FEM; N:PL]	[G:FEM; N:PL]	OK	[G:FEM; N:PL]	[G:FEM; N:PL]	OK

However, things change when first and second person pronouns are concerned, for they do not have morphological specification for gender in either BP or EP. Given that all pronouns but *a gente* are morphologically specified for number in EP and number is the most prominent feature in EP participials, our analysis predicts that in EP participials, all pronouns but *a gente* can value the number feature of participial inflection and be elided. In other words, the pronoun *a gente* in EP should maintain in participial clauses the exceptional behavior it exhibits in finite clauses, because it does not have a value for number. As for BP, our prediction is that syntactic agreement should not be able to license the null counterparts of first and second person pronouns, for the most prominent feature in its participials is gender and these pronouns are not morphologically specified for gender. These predictions are summarized in Table 6.

Table 6. Participial agreement and null subject licensing

Nominative pronouns	EP		BP	
	Syntactic agreement	Null subject licensing via syntactic agreement	Syntactic agreement	Null subject licensing via syntactic agreement
<i>tu</i>	[G:u; N:SG]	OK		
<i>eu</i>	[G:u; N:SG]	OK	[G:u; N:u]	*
<i>você</i>	[G:u; N:SG]	OK	[G:u; N:u]	*
<i>ele</i>	[G:MASC; N:SG]	OK	[G:MASC; N:u]	OK
<i>ela</i>	[G:FEM; N:SG]	OK	[G:FEM; N:u]	OK
<i>a gente</i>	[G:u; N:u]	*	[G:u; N:u]	*
<i>nós</i>	[G:u; N:PL]	OK	[G:u; N:PL]	*
<i>vocês</i>	[G:u; N:PL]	OK	[G:u; N:PL]	*
<i>eles</i>	[G:MASC; N:PL]	OK	[G:MASC; N:PL]	OK
<i>elas</i>	[G:FEM; N:PL]	OK	[G:FEM; N:PL]	OK

At first sight, the data in (19) below show that the above predictions are only partially fulfilled.

- (19) a. [*A gente*] *estava muito confiante*. [*Uma vez* \emptyset]
 we were very confident one time
eleito/eleita], a crise ainda *poderia ser revertida*.
 elect.PPLE.MASC.SG/FEM.SG the crisis still could be reversed
 “We were very confident. Once we were elected, the crisis could still be reversed.” ($\emptyset = a\ gente \rightarrow EP/BP: *$)
- b. [*Depois de* \emptyset *contratado/contratada*], *os meus problemas*
 after of hire.PPLE.MASC.SG /FEM.SG the my problems
acabaram. ($\emptyset = eu \rightarrow EP: OK; BP: ?$)
 finished
 “After I was hired, my problems were over.”
- c. *Não se preocupe*. [*Uma vez* \emptyset *chamada para a*
 not yourself worry one time call.PPLE.FEM.SG for the
entrevista], *as chances de você ser contratada são boas*.
 interview the chances of you.SG be hired are good
 “Don’t worry. Once you have been called for the interview, your chances of being hired are good.” ($\emptyset = você \rightarrow EP: OK; BP: ?$)
- d. *Vocês não perceberam o óbvio*. [*Depois de* \emptyset]
 you.PL not perceived the obvious after of
demitidos/demitidas], *tudo ia ficar mais difícil*.
 fire.PPLE.MASC.PL/FEM.PL everything went stay more difficult
 “You(PL) didn’t see the obvious thing. After you were fired, everything was going to become much harder.” ($\emptyset = vocês \rightarrow EP: OK; BP: ?$)

As predicted, in (19) all pronouns but *a gente* can be null in EP and deletion of *a gente* in BP is not permitted either. The null counterparts of first and second person pronouns in BP contrast with the null counterparts of third person pronouns (see (16)), also as predicted. However, the contrast is not as strong as predicted, for (19b–d) are marginally allowed in BP. Our conjecture is that the semantic agreement independently available for first and second person pronouns is what is behind the marginal licensing of null subjects in (19b–d) in BP. Evidence that this is the case is that (full) semantic agreement can rescue even the null counterpart of *a gente*, as shown in (20) (cf. (19a)).

- (20) [A gente] estava muito confiante. [Uma vez \emptyset eleitos/eleitais],
 we was very confident one time elect.PPLE.MASC.SG/FEM.SG
a crise ainda poderia ser revertida. (EP: OK; BP: ?)
 the crisis still could be reversed
 “We were very confident. Once we were elected, the crisis could still be reversed.”

4. Null subjects without ϕ -agreement: Gerund clauses

Ordinary gerunds in EP and BP can license null subjects in noncontrolled configurations, despite the fact that they are not morphologically specified for person, number, or gender. Under our analysis, null subjects can be licensed in such environments if Case is being computed as the most prominent feature, in consonance with the Prominent Feature Valuation Condition.⁸ That being so, we predict that subject ellipsis in ‘agreementless’, Case-marking gerunds should be possible for all pronouns in both EP and BP. The data in (21) below show that this prediction holds true. BP shows no restrictions as to which pronouns can be null and the null counterpart of *a gente* in EP is no longer exceptional (see (21e)), as its morphological deficiency is in person and number, not Case.⁹

8. This may suggest that Case valuation is not always a reflex of ϕ -agreement (contra Chomsky 2001). In the clauses under discussion, it might be argued that the relevant Case is assigned by default. We leave further discussion of this issue to another occasion.

9. English also has null subjects in gerund clauses, but these are instances of controlled (or arbitrary) PRO, not *pro* (cf. Pires 2006 for discussion and references). Thus, a direct translation of sentences like the ones in (21) into English yields ungrammatical sentences, as shown in (i). Moreover, gerund clauses like (ii), with a null non-referential (quasi-argumental) subject, are also excluded in English, as shown in (iii).

- (i) a. * \emptyset_i entering into the sea, my_i watch stopped working.
 b. * \emptyset_i doing things honestly, everyone is going to vote for you_i.
 c. * \emptyset_i working well, the director will hire {her/him/them}_i next month.
 d. * \emptyset_i working well, the director is going to raise our_i salary this month.
 e. * \emptyset_i eating all the soup, Mom will let us_i go play.

- (21) a. [\emptyset *entrando no mar*], *o meu relógio parou*
 entering in.the sea the my watch stopped
 “When I entered into the sea, my watch stopped working.”
 ($\emptyset = eu \rightarrow$ EP/BP: OK)
- b. [\emptyset *agindo honestamente*], *todos vão votar em você(s)*.
 acting honestly all go vote in you
 “If you do things honestly, everyone is going to vote for you.”
 ($\emptyset = você(s) \rightarrow$ (EP/BP: OK)
- c. [\emptyset_i *trabalhando bem*], *o diretor*
 working well the director
contrata-o_i/-a_i/-os_i/-as_i *no mês que vem.*
 hire-him/-her/-them.MASC/-them.FEM in.the month that comes
 “If {he/she/they} work(s) well, the director will hire him/her/them} next
 month.” ($\emptyset = ele/ela/eles/elas \rightarrow$ EP: OK)
- c'. [\emptyset_i *trabalhando bem*], *o diretor contrata ele_i/*
 working well the director hire him/
ela_i/eles_i/elas_i *no mês que vem.*
 her/them.MASC/them.FEM in.the month that comes
 “If {he/she/they} work(s) well, the director will hire {him/her/them} next
 month.” ($\emptyset = ele/ela/eles/elas \rightarrow$ BP: OK)
- d. [\emptyset *trabalhando bem*], *o diretor vai aumentar o nosso salário*
 working well the director goes raise the our salary
este mês. ($\emptyset = nós \rightarrow$ EP/BP: OK)
 this month
 “If we work well, the director is going to raise our salary this month.”
- e. [\emptyset *comendo a sopa toda*], *a mãe deixa a gente ir brincar.*
 eating the soup all the mother lets us go play
 “If we eat the all soup, Mom will let us go play.”
 ($\emptyset = a gente \rightarrow$ EP/BP: OK)

-
- (ii) *Chovendo, aparecem cogumelos.*
 raining sprout mushrooms
 “When/If it rains, mushrooms sprout.”
- (iii) *Raining, mushrooms sprout.

5. Concluding remarks

It is commonly assumed in the literature that null subjects are licensed by unambiguous morphology. We have argued throughout this paper that rather than overt morphological distinctions, what is relevant for null subject licensing is the underlying feature specification of the verbal inflection, after subject-agreement between T and *pro* values the relevant person/number/Case feature (see Rosenqvist 2018 for relevant discussion). Hence, only close inspection of the pronominal and agreement systems of individual NSLs permits an adequate characterization of them.¹⁰

From this perspective, we expect a lot of variation to be found among partial NSLs, which is empirically confirmed (see Lobo & Martins 2017, for references), and this is compatible with Holmberg's tripartite typology. But the characterization of each type of NSL in Holmberg's typology does not account for the fact that a 'consistent' NSL such as EP may not license ellipsis of a given pronoun. Nor does it capture the fact that ϕ -deficient domains such as participials and gerunds may license null pronouns that are not licensed in finite domains (even in 'consistent' NSLs). Third person pronouns in BP, for instance, have their null counterparts licensed in participials and gerunds, but not in finite clauses. Similarly, in both EP and BP the null counterpart of *a gente* is syntactically licensed in gerunds but not in finite or participial domains. In other words, *a gente* behaves similarly in BP and EP across these three domains, despite the fact that BP is taken to be a partial NSL and EP, a consistent NSL. As a matter of fact, the notion of 'rich agreement' is generally taken to play a role only in finite domains, blurring the distinction between consistent and partial NSLs in nonfinite domains.¹¹

10. One reviewer raises the issue of learnability "if overt morphological distinctions are not particularly relevant". The point is well taken. Nevertheless, it should be noted that children independently have to solve apparent mismatches between the semantic specification of subject pronouns and the morphological specification of the corresponding agreement inflection. So, it does not seem implausible that they use this information to determine the properties of null subjects in their language. Whether this speculation is on the right track and what specific cues children may rely upon to acquire null subjects in EP and BP (with all the subtleties of the latter) is a research topic that goes far beyond the goals of the current paper, though.

11. Note that these intricate facts also present evidence against approaches that take verbal agreement inflection to be pronominal in NSLs (e.g., Barbosa 1995; Alexiadou & Anagnostopoulou 1998; and Kato 1999). Under the pronominal T account, the ϕ -features on T are interpretable, hence not dependent on agreement for valuation. This predicts a much more uniform behavior for the verbal agreement paradigm than what we have seen in EP and BP. We thank a reviewer for calling our attention to this point.

Facts like those discussed in §§ 2–4, coupled with the availability of a generic third person singular indefinite in EP mentioned in § 1, thus raise serious questions about the adequacy of Holmberg’s typology in distinguishing consistent and partial NSLs in terms of the presence of a D-feature on T. The discussion above shows that the same language may behave as a ‘partial’, ‘consistent’, or ‘radical’ NSL depending on the morphological feature specification of its nominative pronouns and T heads.

Acknowledgements

Earlier versions of this article have been presented at Going Romance 2018 (Utrecht University), the workshop *Linguística formal I* (Universidade Federal de Santa Catarina), and the University of Connecticut. We would like to thank the audiences at these presentations, as well as an anonymous reviewer for their valuable comments and suggestions.

Funding

Writing of the current version has been partially supported by FCT – *Fundação para a Ciência e a Tecnologia* (funding UIDB/00214/2020; first author) and CNPq (grant 303195/2019-3; second author), to which we are also thankful.

References

- Alexiadou, Artemis & Elena Anagnostopoulou. 1998. Parametrizing AGR: Word order, V-movement and EPP-checking. *Natural Language and Linguistic Theory* 16. 491–539. <https://doi.org/10.1023/A:1006090432389>
- Barbosa, Pilar. 1995. Null subjects. Ph.D. Dissertation: Massachusetts Institute of Technology.
- Barbosa, Pilar. 2019. *pro* as a minimal *nP*: Toward a unified approach to *pro*-drop. *Linguistic Inquiry* 50. 487–526. https://doi.org/10.1162/ling_a_00312
- Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan. 2009. *Parametric variation: Null subjects in minimalist theory*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511770784>
- Bonet, Eulàlia. 1991. Morphology after syntax: Pronominal clitics in Romance. Ph.D. Dissertation: Massachusetts Institute of Technology.
- Carvalho, Janayna. 2019. Teasing apart 3rd person null subjects in Brazilian Portuguese. In Ingo Feldhausen, Martin Elsig, Imme Kuchenbrandt & Mareike Neuhaus (eds.), *Romance languages and linguistic theory 2016. Selected papers from ‘Going Romance’ 30 Frankfurt*, 237–254. Amsterdam: John Benjamins. <https://doi.org/10.1075/rllt.15.12car>
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.

- Chomsky, Noam. 2008. On Phases. In Robert Freidin, Carlos P. Otero, and María Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory: Essays in honor of Jean-Roger Vergnaud*, 133–166. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9780262062787.003.0007>
- Costa, João & Sandra Pereira. 2005. Phases and autonomous features: A case of mixed agreement in European Portuguese. In Martha McGinnis & Norvin Richards (eds.), *Perspectives on Phases, MIT Working Papers in Linguistics* 49. 115–124.
- Costa, João & Sandra Pereira. 2013. *A gente*: Pronominal status and agreement revisited. *Linguistic Review* 30. 161–184. <https://doi.org/10.1515/tlr-2013-0006>
- Duarte, Maria Eugênia. 1993. Do pronome nulo ao pronome pleno: a trajetória do sujeito no português do Brasil. In Ian Roberts & Mary A. Kato (eds.), *Português Brasileiro: uma viagem diacrônica*, 107–128. Campinas: Editora da UNICAMP.
- Duarte, Maria Eugênia. 2000. The loss of the Avoid Pronoun principle in Brazilian Portuguese. In Mary A. Kato & Esmeralda Negrão (eds.), *Brazilian Portuguese and the Null Subject parameter*, 17–36. Frankfurt: Vervuert-Iberoamericana. <https://doi.org/10.31819/9783964561497-002>
- Duguine, Maia. 2013. Null arguments and linguistic variation: A minimalist analysis of pro-drop. Ph.D. Dissertation: University of the Basque Country/University of Nantes.
- Ferreira, Marcelo. 2000. Argumentos Nulos em Português Brasileiro. M.A. Thesis: Universidade Estadual de Campinas.
- Ferreira, Marcelo. 2009. Null subjects and finite control in Brazilian Portuguese. In Jairo Nunes (ed.), *Minimalist essays on Brazilian Portuguese Syntax*, 17–49. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.142.04fer>
- Galves, Charlotte. 1993. O enfraquecimento da concordância no português brasileiro. In Ian Roberts & Mary A. Kato (eds.), *Português Brasileiro: uma viagem diacrônica*, 387–408. Campinas: Editora da UNICAMP.
- Holmberg, Anders. 2005. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36. 533–64. <https://doi.org/10.1162/002438905774464322>
- Holmberg, Anders. 2009a. Null subject parameters. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.) *Parametric variation: Null subjects in minimalist theory*, 88–124. Cambridge: Cambridge University Press.
- Holmberg, Anders. 2009b. The null generic subject pronoun in Finnish: A case of incorporation. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 200–230. Cambridge: Cambridge University Press.
- Holmberg, Anders. 2016. Null subjects in Finnish and the typology of pro-drop. To appear in Anne Tamm & Anne Vainikka (eds.), *Uralic syntax*. <https://ling.auf.net/lingbuzz/003138>
- Holmberg, Anders, Aarti Nayudu & Michelle Sheehan. 2009. Three partial null-subject languages: A comparison of Brazilian Portuguese, Finnish, and Marathi. *Studia Linguistica* 63. 59–97. <https://doi.org/10.1111/j.1467-9582.2008.01154.x>
- Kato, Mary A. 1999. Strong pronouns, weak pronominals and the null subject parameter. *Probus* 11. 1–37. <https://doi.org/10.1515/prbs.1999.11.1>
- Lobo, Maria & Ana Maria Martins. 2017. Subjects. In Andreas Dufter & Elisabeth Stark (eds.), *Manual of Romance morphosyntax and syntax*, 27–88. Berlin: De Gruyter. <https://doi.org/10.1515/9783110377088-002>
- Menuzzi, Sérgio. 2000. First person plural anaphora in Brazilian Portuguese: Chains and constraint interaction in binding. In João Costa (ed.), *Portuguese syntax: New comparative studies*, 191–240. Oxford: Oxford University Press.
- Merchant, Jason. 2001. *The syntax of silence*. Oxford: Oxford University Press.

- Nunes, Jairo. 2008. Inherent Case as a licensing condition for A-movement: The case of hyper-raising constructions in Brazilian Portuguese. *Journal of Portuguese Linguistics* 7. 83–108. <https://doi.org/10.5334/jpl.129>
- Nunes, Jairo. 2019. Remarks on finite control and hyper-raising in Brazilian Portuguese. *Journal of Portuguese Linguistics* 18(1). 4. <https://doi.org/10.5334/jpl.196>
- Perlmutter, David. 1971. *Deep and surface constraints in syntax*. New York: Holt, Rinehart & Winston.
- Pires, Acrísio. 2006. *The minimalist syntax of defective domains. Gerunds and infinitives*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.98>
- Roberts, Ian. 2010. A deletion analysis of null subjects. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 58–87. Cambridge: Cambridge University Press.
- Rosenkvist, Henrik. 2018. Null subjects and distinct Agreement in Modern Germanic. In Federica Cognola & Jan Casalicchio (eds.), *Null subjects in Generative grammar*, 285–306. Oxford: Oxford University Press.
- Saab, Andrés. 2008. Hacia una teoría de la identidad parcial en la elipsis. Ph.D. Dissertation: University of Buenos Aires.
- Saab, Andrés. 2016. On the notion of partial (non-) *pro*-drop in Romance. In Mary A. Kato & Francisco Ordoñez (eds.), *The morphosyntax of Spanish and Portuguese in Latin America*, 49–77. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190465889.003.0003>
- Sheehan, Michelle. 2006. The EPP and null subjects in Romance. Ph.D. Dissertation: Newcastle University.
- Sheehan, Michelle. 2016. Subjects, null subjects and expletives. In Susann Fischer & Christoph Gabriel (eds.), *Manual of grammatical interfaces in Romance*, 329–362. Berlin: De Gruyter. <https://doi.org/10.1515/9783110311860-014>
- Simioni, Leonor. 2011. Concordância em construções passivas com argumentos pré- e pós-verbais e incorporação de nomes nus no Português Brasileiro. Ph.D. Dissertation: Universidade de São Paulo.
- Wurmbrand, Susi. 2014. The Merge condition: A syntactic approach to selection. In Peter Kosta, Steven Franks, Teodora Radeva-Bork & Lilia Schürcks (eds.), *Minimalism and beyond: Radicalizing the interfaces*, 130–166. Amsterdam: John Benjamins. <https://doi.org/10.1075/lfab.11.06wur>

Aspect in the acquisition of the Spanish locative paradigm by Italian L2 learners

Sílvia Perpiñán¹ and Rafael Marín^{2,3}

¹Universitat Pompeu Fabra / ²CNRS / ³Université de Lille

The present study investigates the development of the expression of the locative paradigm in the L2 Spanish of Italian-speaking learners. We investigate (i) whether the developmental stages proposed for English-speaking learners (VanPatten 1987; Perpiñán, Marín & Moreno Villamar 2020) hold for Italian-speaking learners; and (ii) whether Italian, a language that partially overlaps with the distribution of the Spanish copulas has a facilitative role in the process. 33 Italian-speaking learners of Spanish and 21 monolingual Spanish speakers completed a short proficiency test, an acceptability judgement task, and a picture matching task targeting these constructions. Results indicate that unlike what VanPatten (1987, 2010) has proposed for English-speaking learners of Spanish, Italian speakers do not present a delay in the acquisition of *estar*, but instead, it is overproduced in locative contexts from very early on. We argue that this overproduction of *estar* is due to the readily available mapping of ‘temporal boundedness’ with *estar* in the grammar of these L2 learners, whereas the presence of the feature ‘dynamicity’, even though it is relevant in the distribution of copulas in Italian, comes later in L2 development.

Keywords: copulas, *ser*, *estar*, temporal boundedness, dynamicity, aspect, locative, developmental stages, L2 Spanish, L1 Italian

1. Introduction

It is at the core of generative linguistics to describe what is universal about language and what varies. As a result, one of the main questions in generative second language research is whether the same processes and developmental paths apply to all types of second language learners’ interlanguages. This study provides novel data that can speak to the universality (or lack thereof) of transitional stages during the acquisition of a second language. In particular, we investigate the developmental path in the acquisition of *ser* and *estar* in the Spanish interlanguage of Italian

speakers, an understudied language combination. We explore the selection of the copulas *ser* and *estar* in locative predicates in both, location of an individual, as in *Juan está en la cocina* (“Juan is in the kitchen”), and location of an event, as in *La fiesta es en la playa* (“The party is at the beach”). Both of these uses are expressed in Italian with the copula *essere*, whereas Spanish selects *ser* or *estar* depending on the aspectual nature of the subject, in particular whether it is [\pm dynamic] (Perpiñán et al. 2020).

Adopting a feature analysis (Chomsky 1995, 2001), we assume that semantic features may appear in different configurations in the languages at play; when acquiring an L2, the learner needs to reassemble the “features from the way these are represented in the first language into new formal configurations on possibly quite different types of lexical items in the L2” (Lardiere 2009: 173). This theoretical proposal implies full transfer of the L1 and at the same time allows room for restructuring the interlanguage grammar, most likely guided by UG – ideas that directly stem from the Full Transfer / Full Access Hypothesis (Schwartz & Sprouse 1996). Hence, the task of the L2 learners is to detect, select and (re)assemble, if needed, the appropriate semantic features in the specific functional lexicon.

VanPatten (1987, 2010) identified, as the main difficulty in the acquisition of the Spanish copulas, a marked aspectual property of *estar*, which would be responsible for its delayed acquisition. VanPatten proposed that L2 learners start out with *ser* and later acquire *estar*. However, when describing this developmental path, he only took into account data from English-speaking learners of Spanish, which may not be generalizable to other learning contexts. Indeed, one of our main research questions is whether the developmental path described in VanPatten (1987, 2010) for (mostly) English-speaking L2 learners of Spanish would also hold for Italian-speaking learners of Spanish, or whether language-specific effects such as L1 transfer – either negative or positive – would surface in the acquisition process, changing the described developmental path. We also question if, in the course of reassembling the specific feature bundles and mapping them onto the L2 functional lexicon, there will be semantic features that are mapped or properly acquired faster than others.

In order to answer this, we first explore the microparametric differences that characterize Spanish and Italian (and Southern Italian) in the expression of locative predicates with *ser/essere* and *estar/stare*, and then describe and compare the developmental path of these L2 learners with that of English-speaking learners, as described by VanPatten (1987) and, more recently, by Perpiñán et al. (2020).

2. The linguistic phenomenon: Locatives with copulas

In Spanish, individuals are located by *estar* (1a), while events are located with *ser* (1b).

- (1) (Spanish)
 - a. *María está/*es en casa*
“María is at home.”
 - b. *La fiesta es/*está en la playa*
“The party is at the beach.”

In Standard Italian, only *essere* is used to locate both individuals (2a) and events (2b). *Stare* can also be used with individuals, but the meaning conveyed does not include location alone: an additional nuance of remanence, staying in a place, is also introduced (3).

- (2) (Standard Italian)
 - a. *Maria è/#sta a casa*
“Maria is at home.”
 - b. *La festa è/*sta in spiaggia*
“The party is at the beach.”
- (3) (Standard Italian)
 - a. *Maria sta a casa questa sera*
“Maria is staying at home tonight.” (meaning = she won’t go out tonight)
 - b. *Le chiavi stanno sopra la mensola*
“The keys are usually kept on the shelf.”

Since our data were collected in Rome and our speakers were potentially exposed to Romanesco, the vernacular variant of Italian spoken in Rome, we describe it as well, although the L2 speakers tested in this study declared Italian and not *dialetto* to be their mother tongue (for more details see the participants section). In Romanesco, we observe essentially the same pattern as in Spanish: individuals are located with *stare* (4a), while events are located with *ser* (4b).

- (4) Romanesco
 - a. *Maria sta/*è a casa*
“Maria is at home.”
 - b. *La festa è/*sta in spiaggia*
“The party is at the beach.”

Indeed, Serianni & Castelvocchi’s (1991) Italian Grammar indicates that *stare* used as *essere* is considered a Southern variant (*un meridionalismo*) in sentences such as *mio zio sta al circo* “my uncle is at the circus” (example from Siniscalchi 1889: 143, cited in Serianni & Castelvocchi 1991: 436). Thus, in Southern Italian, *stare*

and *essere* seem to share the same distribution as in Spanish, whereas in Standard Italian, only *essere* is a true copula, and *stare* is a pseudo-copula that obligatorily includes the idea of remanence when used for locative purposes. In any case, if our participants had been exposed directly or indirectly to Romanesco, they would have both the Standard Italian and *dialetto* systems available (see footnote 2)

The fact that *stare* is not a ‘real’ copula in Standard Italian, and thus not in direct opposition with *essere*, could explain the lack of theoretical or applied studies that compare these two verbs in Italian. The obvious similarities and yet crucial morphosyntactic and semantic differences in the distribution of copulas in Spanish and Southern Italian, on the one hand, and Standard Italian on the other, makes this structure an ideal locus of investigation for testing the role of previous linguistic experience and transfer effects in the acquisition of Spanish copulas, particularly in languages that are very closely related. Table 1 summarizes the distribution of copulas in Spanish and Standard Italian.

Table 1. Distribution of copulas to express location

	Spanish		Standard Italian	
	<i>ser</i>	<i>estar</i>	<i>essere</i>	<i>stare</i>
Events	+	–	+	–
Individuals	–	+	+	–*

* In Standard Italian, *stare* is used with individuals to express remaining in a location.

The dichotomy in location with *ser/estar* in Spanish is so robust that there is a type of noun that can have a double reading depending on their selection of *ser* or *estar*. That is, Spanish lexicalizes certain semantic properties of the subject: the copula varies depending on its combination with an object or an event. If the ambiguous noun selects *ser*, then it must refer to an event (5a), but if it selects *estar*, it must refer to an object (5b).

- (5) a. *Los fuegos artificiales son en el parque.*
 “The fireworks are (= will happen) in the park.”
 b. *Los fuegos artificiales están en el parque.*
 “The fireworks (= the dynamite) are in the park.”

In Standard Italian, these two meanings of *fireworks*, the event reading and the object reading, would combine with *essere*, rendering it ambiguous. If it combined with *stare*, it would mean that the fireworks need to be in, or could not leave, the park.

The series of matrices of features summarized in (6) shows that Italian and Spanish are in an asymmetrical superset/subset relationship with respect to locative

constructions. Italian *essere* can be used for both types of subjects, individuals [–dynamic] and events [+dynamic], whereas Italian *stare* is more restricted than Spanish *estar*.¹

- (6) a. **Italian**
essere
- [±dynamic
+temporally bounded]
- stare*
- [–dynamic
+temporally bounded
+remain]
- b. **Spanish**
ser
- [+dynamic
+temporally bounded]
- estar*
- [–dynamic
+temporally bounded]

By means of the [±dynamic] feature we want to capture the fact that in Standard Italian both individuals [–dynamic] and events [+dynamic] are located with *ser*. On the other hand, in order to capture that *stare* is used in Italian only to express permanence (of an individual) in a place, an additional feature [+remain] is postulated.

3. Research questions and learnability tasks

Our overarching research question refers to the developmental path that Italian-speakers would follow to acquire the Spanish copulas in locative contexts. In particular, we want to know whether the developmental path already described for English-speaking learners (VanPatten 1987; Perpiñán 2014; Perpiñán et al. 2020) would also apply for Italian-speaking learners, or whether it is language specific. Specifically, (i) we want to examine the initial stages of acquisition, and in particular

1. If we limit our analysis to the expression of location, the feature [temporally bounded], always marked positively in (6), might be perceived as a rather redundant feature, as all location is inherently temporally bounded. However, this specification is crucial if we want to provide an analysis compatible with a general account of the acquisition of Spanish copulas, given that cases such as *ser* + IL adjective (*Juan es budista* ‘Juan is Buddhist’), denote temporally unbounded predicates (Arche, Fábregas & Marín 2017; Perpiñán et al. 2020).

the role of transfer and/or universal processes. Given the obvious similarities between the two languages, we question whether L1-Italian learners transfer their Italian locative system into Spanish at the initial stages of acquisition. (ii) If so, how do they recover from L1-transfer and acquire the target distribution and interpretation of Spanish locative constructions? In which order does feature reassembly occur? Are some features fully acquired before others? (iii) And finally, what is more difficult: to move from a subset grammar to a superset grammar (broaden the uses of *stare*) or to move from a superset grammar to a subset grammar (reduce the uses of *ser*)?

In light of the microparametric differences and similarities between Standard Italian and Spanish, and if Full Transfer applies, then we will find that the starting point for Italian L2 learners is to locate with *essere* = *ser*, which is less specified than *stare*. As a result, we hypothesize that at the initial stages of acquisition, Italian-speaking L2 learners would generally locate both objects and events with *ser*; *stare* would be reserved for a small subset of locative predicates: those that imply [+remanence], which unfortunately are not considered in the experimental materials. If the L2 learners start out with this Italian-looking interlanguage grammar, then their learnability task is to: (i) broaden their uses of *estar* to all temporally bounded locations, which implies moving from a subset grammar (Italian *stare* which only applies to ‘remained location’) to a superset grammar (Spanish *estar*, which is used for the location of all individuals); and (ii) reduce their use of *ser* to locate only [+dynamic] subjects, that is, to move from a superset grammar (Italian *essere* which allows all sorts of locations), to a subset grammar (Spanish *ser*, which can only locate events). In short, they need to widen their *estar* grammar by unselecting or unlearning the ‘remanence’ feature specification, and they need to reduce their *ser* grammar by selecting the new ‘dynamicity’ requirement. This requires a non-trivial reassembly of features onto an almost identical functional lexicon. Several studies have proposed that reassembly of features in an existing category is harder than selecting features for a new category (Ionin & Montrul 2010; Sprouse 2006). This might be because reassembly usually requires unlearning or deselecting a feature, which cannot be done through positive evidence alone, and negative evidence, unless taught explicitly, is rarely present.

4. Our study

In order to address the research questions laid out in § 3, we conducted a small-scale study with adult L1-Italian speakers, learning Spanish in Rome (Italy), of different proficiency levels, so we can cross-sectionally outline the developmental path of this target structure.

4.1 Participants

Participants included an experimental group of 33 Italian-speaking learners of Spanish, mean age at testing 34.3; and a control group of 21 monolingual Spanish native speakers, mean age at testing 27.6. The Spanish learners were all studying or had studied Spanish in the Instituto Cervantes of Rome, for a period of time ranging from 3 months to 4 years, and most of them were university students who had already graduated from college. These participants were further divided into two groups (upper-beginners and upper-intermediate) according to their level of proficiency.

4.2 Methods

All participants took a linguistic background questionnaire that asked about the onset of acquisition, language use, courses taken, place of origin, etc. Both groups completed the multiple-choice section (25 items) of the standardized superior DELE (*Diploma de Español para Extranjeros*) test, an Acceptability Judgment Task (AJT), and a Picture Selection Task (PST); all tests were taken in a web survey platform.

The AJT consisted of a total of 110 sentences, randomized and counterbalanced for grammaticality, to be judged using a slider bar on a scale from ‘sounds bad’ to ‘sounds good’. The position of the slider bar later translated into a numeral scale, from 1 (sounds bad) to 100 (sounds good). It targeted 4 conditions with locative structures with *ser* and *estar*, with objects and events. Each condition included 5 tokens ($4 \times 5 = 20$ sentences).

(7) *Ser / Estar with Objects (2 conditions \times 5 tokens = 10)

a. *El libro *es/está encima de la mesa.*

The book *SER / ESTAR on-top of the table

“The book is on the table.”

Ser/ *Estar with Events (2 conditions \times 5 tokens = 10)

b. *La reunión es / *está en el hotel Majestic.*

The meeting SER/ *ESTAR in the hotel Majestic

“The meeting is at hotel Majestic.”

The Picture Selection Task consisted of a sentence and a pair of images; participants had to select the picture that best represented the sentence given. For this task, we used ambiguous nominals, which can be interpreted as either events or objects, depending on the context and the copular verb employed. These ambiguous nominals forced the participants to pay attention to the copular verb used, and to the

interpretation of the subject. There was a total of 12 pairs of pictures, and each pair appeared twice: once targeting an object and once targeting an event. This made a total of 24 scenarios, half of which were controls.

The target scenarios ($k = 12$) consisted of 6 pairs of pictures referring to an ambiguous nominal in its objective or in its eventive reading. These 6 pairs of images appeared each twice: once alongside a sentence with *ser*, thus forcing the eventive reading, and once with *estar*, targeting the object reading. The ambiguous nominals were *la cena* (“the dinner”), *los fuegos artificiales* (“the fireworks”), *la película* (“the movie”), *la obra de Shakespeare* (“Shakespeare’s play”), *el examen final* (“the final exam”), and *el discurso* (“the speech”). All these nouns can refer to a physical object or an event. Figure 1 provides a sample task item targeting the ambiguous nominal *la cena* (“the dinner”) in its objective reading (scenario A), and its eventive reading (scenario B).

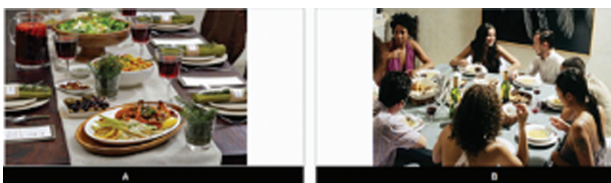


Figure 1. Example of target scenarios in the picture selection task

- Scenario A: *La cena está en el comedor.*
 “The dinner is_{estar} in the dining room.”
 (Subject = object; Target response: Picture A)
- Scenario B: *La cena es en el comedor.*
 “The dinner is_{ser} in the dining room.”
 (Subject = event; Target response: Picture B)

4.3 Results

Linguistic background questionnaire and proficiency test

All L2 learners reported having Italian as their mother tongue. Out of these 33 L2 learners, 22 were born in Rome, 4 outside Rome within the Lazio Province, 3 in Campania, 1 in Puglia, 1 in Tuscany, 1 in Ancona, and 1 in Veneto. Thus, 30 out of 33 are from Central-Southern Italy, and all lived in Rome at the time of testing. When asked about other languages they spoke, including dialects, only 4 participants chose also ‘dialect’, and ranked it 4th, 5th, and 7th among languages they spoke, after Italian, English, or Spanish. These were speakers of Neapolitan, Sabino, Romanesco, and Anconitano. For this reason, and even though we are cautiously

taking into account that these speakers might be potentially exposed indirectly to Romanesco or Southern Italian, our point of departure is Standard Italian, as they self-reported.²

The results from the reduced standardized proficiency test (DELE superior) were converted into accuracy percentages per person, and later clustered into three distinct groups: the control group, and two subgroups within the experimental pool of participants. The Spanish monolingual speakers ($n = 21$) had a mean accuracy of 94% ($SD = 5.5$), the lower proficiency group of the Italian speakers ($n = 14$) had a mean accuracy of 50% ($SD = 8.5$), and the more advanced group ($n = 19$) had a mean accuracy of 75% ($SD = 7$). All three groups were significantly different ($p < .001$). It is important to note that our lower proficiency group is roughly equivalent in proficiency to the low-intermediate English-speaking learners depicted in Perpiñán et al. (2020) –VanPatten’s studies do not provide an independent measure of proficiency. Thus, despite their short exposure to the target language (most of these participants reported to have studied Spanish for a few months, between 2 to 18 months, in an average of 3 hours a week), Italian speakers do not usually count as true beginning learners of Spanish given the similarities between the two Romance languages. Furthermore, the grammar section of the superior DELE used here (version 2005, complete Section 2 of multiple-choice format) does not present a single item in which the dichotomy between *ser/estar* is tested, as most of their items are on collocations, preposition, and subjunctive use. Therefore, it is reasonable to assume that even though these learners might have an overall low-intermediate grammar level as assessed by the DELE, their knowledge of *ser/estar* may still remain incomplete. We call them here ‘upper-beginners’ to reflect better their real experience with the target language.

Acceptability judgment task

The mean averages of the AJT are displayed in Figure 2. In this task, participants had to judge with a slider bar whether the sentences sounded good (100) or bad (1), these mean ratings per condition and person were the dependent variable. The results of a Repeated Measures ANOVA with ‘dynamicity’ (object vs. event), ‘copula’ (*ser* vs. *estar*), and ‘group’ (native vs. upper-beginner vs. upper-intermediate)

2. An anonymous reviewer questions our assumption with respect to the language or variety these speakers have been exposed to at home or in their primary environment. We assume that, for the most part, these speakers have been raised in Italian and not Romanesque, otherwise they would have expressed that in the language background questionnaire. It is still probable, though, that they have heard some uses of *stare* as a plain locative, without the ‘remanence’ component, in the streets of Rome or in informal gatherings. The impact of that overhearing, then, remains for further research by comparing these results with those from similar learners from Northern Italy.

indicated a main effect of ‘copula’ ($F(1, 50) = 25.517, p < 0.001$) since *estar* had overall higher acceptance rates than *ser*; a main effect of ‘group’ ($F(2, 50) = 5.918, p = .005$); but no significant effect of ‘dynamicity’. The between-subjects analysis showed a significant effect of ‘group’ ($F(2, 50) = 5.918, p = .005$); there was a significant interaction of ‘dynamicity’ * ‘copula’ ($F(1, 50) = 124.344, p < .001$), and a significant interaction of ‘copula’ * ‘group’ ($F = 5.655, p = .006$).

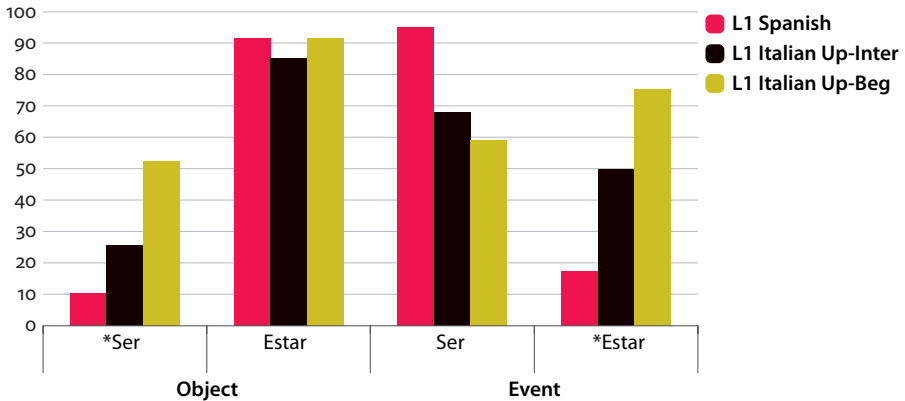


Figure 2. Mean acceptability rates per condition and group of the AJT

To further explore these interactions, we ran an overall post-hoc comparison with Bonferroni correction, which indicated that the more proficient L2 learners’ group did not differ from the control group, but the lower-proficiency learners did ((L1 Spanish = L1 Italian Upper-Intermediate) \neq L1 Italian Upper-Beginners). By conditions, the post-hoc tests found differences between both groups of Italian learners and the native speakers in the conditions with Events (Event + *ser* and Event + *estar*, $ps < .005$), and a significant contrast between the ratings of the monolingual control group and those of the upper-beginner group in the Object + *ser* condition ($p < .001$). There were no differences among the three groups in their acceptance of Objects with *estar*. So, even at relatively early stages of acquisition, Italian speakers know that *estar* localizes objects in Spanish. At the same time, though, both groups of L2 learners accepted the localization of Events with *estar*, which is ungrammatical in all three variants. In this condition (*event + *estar*), all three groups are statistically different from each other ($ps < .005$). A series of Pearson correlations show that only at higher proficiency levels, proficiency correlates with the rejection of *estar* with Events ($r(19) = -.72, p$ (two-tailed) $< .001$), that is, as proficiency increases, the acceptance of *estar* + Event decreases. Interestingly, this correlation is non-significant in the lower-proficiency L2 learners, which means that there is no progression in this structure at the beginning of the acquisition

process: the rejection of *estar* to locate [+dynamic] subjects comes later in the acquisition path. In short, Italian speakers overuse Spanish *estar* to locate both objects and events; the ungrammatical structure with *estar* (Event + *estar*) is accepted even at upper-intermediate levels, unlike the ungrammatical structure with *ser* (Object + *ser*) which becomes more target-like earlier, with no differences between the upper-intermediate L2 learners and the native speakers ($p > .01$).

Picture selection task

The purpose of this comprehension task was to test whether L2 learners are able to distinguish an object from an event just from the choice in copula. Only if the learner has established the association between event and *ser*, on the one hand, and object and *estar*, on the other, they will be able to select the right picture based on the sentence provided. Results are presented in mean proportions of accuracy in Figure 3, where 1 means correct and 0 means incorrect picture selection.

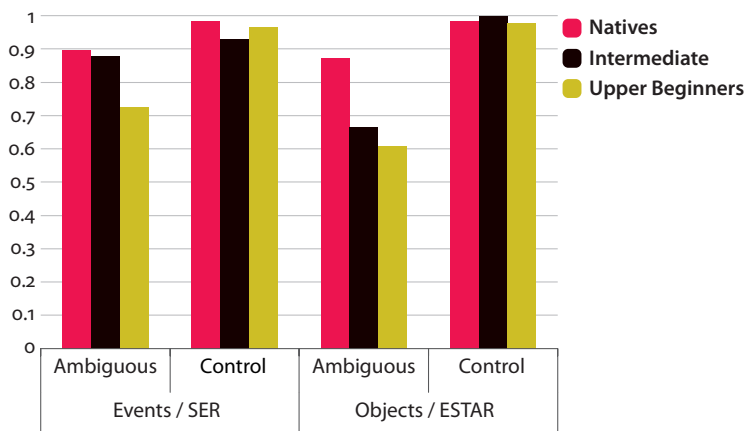


Figure 3. Mean proportions of accurate responses in picture selection task

The mean proportion of accurate responses per condition and group were calculated, Arcsine-transformed, and later submitted to a Repeated Measures ANOVA with ‘Dynamicity’ (event vs. object) and ‘Targetness’ (target structure vs. control structure), and ‘Group’ (upper-beginner, intermediate, native) as independent variables. This revealed a main effect of Dynamicity ($F(1, 51) = 5.447, p = .024$), as events, which are [+dynamic], tend to be appropriately matched with *ser* more often than objects [–dynamic] are matched with *estar*. This result in itself is important, as previous studies have suggested that events are more complex cognitively-speaking, and hence later acquired (Sera 1992). Additionally, even though there is a main effect of ‘Group’ ($F(1, 51) = 5.558, p = .007$), the three groups did not differ in their accuracy

rates with respect to the control structures, those that did not include ambiguous event/object nominals. So, there was a main effect of ‘Targetness’ ($F(1, 51) = 45.225$, $p < .001$), a significant interaction of ‘Dynamicity’ * ‘Targetness’ ($F(1, 51) = 22.308$, $p < .001$), and a three-way interaction of ‘Dynamicity’, ‘Targetness’, and ‘Group’ ($F(2, 51) = 6.358$, $p = .003$).

To further explore these significant interactions, we carried out a series of post-hoc tests with the Bonferroni adjustment. Within the target conditions – the only ones with differences – both L2 groups significantly differed from the control group in their accuracy of object selection when given a sentence with *estar* ($ps < .05$); with respect to the selection of a picture referring to an event when *ser* was provided, the more advanced group had accuracy rates comparable to those of the control group ($p = 1$), but the upper-beginner group differed from the monolingual group ($p < .05$).

5. Discussion and conclusion

Our results largely indicate that Italian speakers generally do not have a delay in the acquisition of Spanish *estar*, as has been proposed for English-speaking learners (VanPatten 2010) and also for Arabic speakers (Perpiñán 2014). The L2 learners do not differ from the monolingual native speakers in their acceptance rates of object + *estar* constructions, but they overextend *estar* to ungrammatical cases (with events). Thus, they quickly restructure their grammars associating *estar* with all [+temporally bounded] predicates, even with those that do not combine with *estar* – the events. The result is that the upper beginner learners over-accept all locatives with *estar*. We conclude then, that moving from a smaller, more restricting grammar, that of Italian *stare*, to a bigger one, that of Spanish *estar*, does not initially pose difficulties. The problem appears when the L2 learners need to constrain this *estar* over-expanded grammar. In particular, these learners have difficulties realizing that ‘dynamicity’ is a relevant feature that has grammatical consequences for the expression of location in Spanish. Interestingly, this overextension of *estar* cannot be explained by L1 transfer from neither variant, Standard Italian nor Romanesco, and it is also found to a lesser extent in the interlanguage grammar of English-speaking learners of Spanish (Perpiñán et al. 2020). Thus, overextension of *estar* reflects a typical L2 interlanguage stage that is found systematically in the acquisition of the Spanish locative paradigm regardless of the L1. Finally, we conclude that, as already found in other studies (Ionin & Montrul 2010), unlearning or restricting a grammar is more difficult than expanding it.

With respect to *ser*, both L2 learner groups are localizing / identifying events with *ser* relatively well, but the beginner group is still accepting object location with *ser* to a significant extent, which is ungrammatical. The paradoxical fact that the upper-beginner acceptance rates of *ser* with events are more target-like than those of the upper-intermediate group makes us interpret the beginners' results of *ser* with events (ungrammatical) and objects (grammatical) as an initial across-the-board extension of the more basic, less marked copula: *ser*. The intermediate group's use of *ser*, on the other hand, distinguishes locating objects reasonably well from locating events. Indeed, their accuracy responses in the Picture Selection Task were target-like when they were provided with a sentence with *ser* and had to select an event. This was not the case with *estar* sentences that required the selection of an image depicting an object, which were harder for all L2 learners.

We propose, then, an initial short interlanguage stage in which *ser* is used to locate all [+temporally bounded] predicates; this stage can be due to either direct L1 transfer from Italian or to the selection of the less aspectually specified copula: *ser*. Unfortunately, our experimental design does not allow us to tease apart these two causes. Italian L2 learners soon hypothesize that Spanish generally localizes with *estar*, resulting in the partially ungrammatical association of all [+temporally bounded] locative predicates with *estar*; even those that require *ser* (events). This would constitute the second developmental stage in which the feature 'dynamicity' is still unspecified or simply not yet selected. Both of these stages produce ungrammaticalities for overextension, which are difficult to recover from, particularly the long-lasting association *estar* = location ([+ temporally bounded predicates]). The 'dynamicity' feature is not selected or specified until later in the interlanguage grammar of these L2 speakers.

Returning to our research questions, we conclude that L1 transfer plays a moderate role at the very beginning of the development, and that L2 learners recover relatively fast from this initial stage. The stage in which Italian speakers generally locate with *estar* appears very soon in the developmental path, maybe favored by the exposure to Southern varieties, but we cannot ascertain that Romanesco directly transfers onto the interlanguage grammar of these speakers. In fact, English-speaking learners go through the same developmental stages (Perpiñán et al. 2020), but at a different pace. To sum up, the real difficulty in the acquisition process consists of restricting the superset grammars because, in the case at hand, this requires selecting a new, not-always-evident aspectual feature: 'dynamicity' of the subject.

References

- Arche, María J., Antonio Fábregas & Rafael Marín. 2017. Towards a unified treatment of Spanish copulas. In Silvia Perpiñán, David Heap, Itziri Moreno-Villamar & Adriana Soto-Corominas (eds.), *Romance languages and linguistic theory 11. Selected papers from the 44th Linguistic Symposium on Romance Languages (LSRL)*, 33–52. Amsterdam: John Benjamins.
<https://doi.org/10.1075/rllt.11>
- Chomsky, Noam. 1995. Minimalist inquiries: The framework. *MIT Occasional Papers in Linguistics* 15. 1–56.
- Chomsky, Noam. 2001. Derivation by Phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge: MIT Press.
- Ionin, Tania & Silvina Montrul. 2010. The role of L1 transfer in the interpretation of articles with definite plurals in L2 English. *Language Learning* 60(4). 877–925.
<https://doi.org/10.1111/j.1467-9922.2010.00577.x>
- Lardiere, Donna. 2009. Some thoughts on the contrastive analysis of features in second language acquisition. *Second Language Research* 25(2). 173–227.
<https://doi.org/10.1177/0267658308100283>
- Perpiñán, Silvia. 2014. Locatives and existentials in L2 Spanish: The acquisition of the semantic contrasts among *ser*, *estar* and *haber*. *Second Language Research* 30(4). 485–513.
<https://doi.org/10.1177/0267658313518215>
- Perpiñán, Silvia, Rafael Marín & Itziri Moreno Villamar. 2020. The role of aspect in the acquisition of *ser* and *estar* in locative contexts by English-speaking learners of Spanish. *Language Acquisition* 27(1). 35–67. <https://doi.org/10.1080/10489223.2019.1610408>
- Schwartz, Bonnie D. & Rex A. Sprouse. 1996. L2 Cognitive states and the Full Transfer/Full Access Model. *Second Language Research* 12(1). 40–72. <https://doi.org/10.1177/026765839601200103>
- Sera, Maria D. 1992. To Be or to Be: Use and acquisition of the Spanish copulas. *Journal of Memory and Language* 31(3). 408–427. [https://doi.org/10.1016/0749-596X\(92\)90021-0](https://doi.org/10.1016/0749-596X(92)90021-0)
- Serianni, Luca & Alberto Castelveccchi. 1991. *Grammatica italiana: Italiano comune e lingua letteraria* (2nd edn.). Turin: UTET.
- Sprouse, Rex A. 2006. Full Transfer and Relexification: Second language acquisition and creole genesis. In Claire Lefebvre, Lydia White & Christine Jourdan (eds.), *L2 acquisition and creole genesis: Dialogues*, 169–181. Amsterdam: John Benjamins.
<https://doi.org/10.1075/lald.42.11spr>
- VanPatten, Bill. 1987. Classroom learners' acquisition of *ser* and *estar*: Accounting for developmental patterns. In Bill VanPatten, Trisha R. Dvorak, & James F. Lee (eds.), *Foreign language teaching: A research perspective* 61–75. Cambridge: Newbury House.
- VanPatten, Bill. 2010. Some verbs are more perfect than others: Why learners have difficulty with *ser* and *estar* and what it means for instruction. *Hispania* 93(1). 29–38.

Catalan nativization patterns in the light of weighted scalar constraints

Clàudia Pons-Moll and Francesc Torres-Tamarit
Universitat de Barcelona / CNRS, Paris 8

In this paper we analyze, from an experimental and formal perspective, the interaction and the implicational relationships between vowel reduction and word-final nasal deletion in Catalan loanwords. We present the results of both a production and a perception test carried out with 31 young speakers from the Barcelona area. Loanwords susceptible to undergoing both nasal deletion and vowel reduction display different patterns, which, according to the tests, show different degrees of likeliness. The most common pattern is underapplication of both processes, followed closely by underapplication of nasal deletion alone and at a large distance by the application of both processes. Finally, underapplication of vowel reduction and application of nasal deletion is unattested in the production test and obtains a very low score in the perception test, that is, it is a very unlikely nativization pattern. The typology of possible nativizations and the implicational relationships between the processes under scrutiny are analyzed in the framework of Harmonic Grammar under Weighted Scalar Constraints, following recent proposals by Hsu & Jesney (2017, 2018).

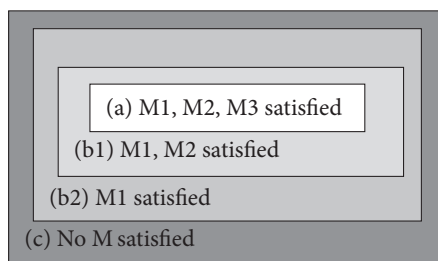
Keywords: loanwords, nativization patterns, implicational relationships, harmonic grammar, weighted scalar constraints, Catalan

1. Introduction

Here we explore the implicational relationships between two or more phonological processes that are susceptible to co-occurring within the same loanword; that is to say, situations in which if a process A applies, so does a process B, but not the other way around. It is well known that loanwords can comply with the markedness constraints satisfied by native words (cf. the Catalan loan *tobogan* “toboggan”, realized as $t[u]b[u]ga[\emptyset]$, with nasal deletion and vowel reduction, as a native word such as *crostó* $cr[u]stó[\emptyset]$ “heel.SING.”; cf. *cr[ɔ]sta* “scab”, *crosto[n]s* “heel.PLUR.”); but usually they comply only with a subset of these markedness constraints (cf. the

same Catalan loan, realized as $t[u]b[u]ga[n]$, with just vowel reduction), and, in many cases, with none (cf. again the same Catalan loan, realized as $t[o]b[o]ga[n]$). It has been argued that this situation reveals a nested core-periphery structure of the lexicon, with three different strata (Itô & Mester 1995, 1999, 2008): (i) ‘the core stratum’, in which loanwords behave like native words and satisfy all markedness constraints, and which contains the nativized loanwords (1a); (ii) ‘the intermediate stratum’, in which loanwords satisfy only a subset of the markedness constraints active in the core stratum, and which includes the partially nativized loanwords (1b1, 1b2), and (iii) ‘the peripheral stratum’, in which loanwords do not satisfy any of the markedness constraints active in the previous strata, and which comprises the non-nativized loanwords (1c).

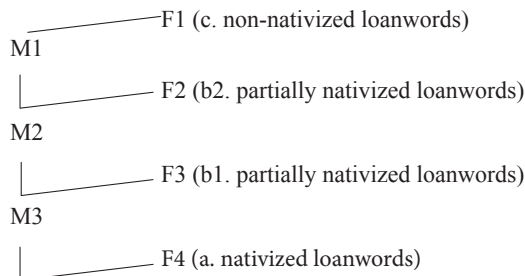
- (1) Core-periphery structure of the lexicon (Itô & Mester 1995, 1999, 2008)



In Itô & Mester’s model, the differences between these strata are explained by the variable position of a block of faithfulness constraints $F1, F2, F3 \dots$, to which lexical items in each stratum are indexed, with respect to a language-particular fixed hierarchy of markedness constraints ($M1 \gg M2 \gg M3$). As illustrated in (2), in the nuclear stratum, all markedness constraints relevant to explaining the application of the native processes in a given language outrank faithfulness, and this leads to fully nativized patterns; in contrast, in the peripheral stratum, faithfulness outranks markedness, which leads to non-nativized patterns; finally, in the intermediate stratum, faithfulness is ranked in between the markedness constraints, so that partially nativized patterns – which satisfy both faithfulness and some markedness constraint(s) – are obtained.

As noted by Itô & Mester (2008), a structure of this kind gives rise to asymmetrical implicational patterns in the adaptation of loanwords: “Structures ... are built out of a network of implicational relations involving lexical items and phonological constraints of the following kind: items that are subject to constraint *A* are also always subject to constraint *B*, but not all items subject to *B* are also subject to *A*” (Itô & Mester 2008: 554). For instance, given the structure in (2), items subject to constraint $M3$ are necessarily also subject to constraints $M2$ and $M1$, but items subject to constraint $M2$ are not necessarily subject to constraint $M3$, because faithfulness can intervene in between.

(2) Differences across strata (Itô & Mester 1995, 1999, 2008)



This paper presents the results of two surveys that provide quantitative support for these kinds of interactions in Catalan loanword phonology, and formally accounts for them under the Weighted Scalar Constraints version of Harmonic Grammar, following recent proposals by Hsu & Jesney (2017, 2018).

2. Data

In Catalan loanwords, several processes may interact in the way described in (1) and (2), involving word-final nasal deletion and vowel reduction (which affects low and mid vowels in Eastern Catalan), vowel reduction and vowel laxing of high-mid stressed vowels, and vowel reduction and word-final *-r* deletion. This occurs in loans that contain structures susceptible to undergoing these processes, like *tobogan*, *orangutan*, *xarleston*, *wonton*; *pòster*, *mòdem*, *euro*, *rècord*; *sommelier*, *atelier*, *dossier*, *necesser* ‘toilet bag’, respectively. In this paper we deal with the first type of interaction, the one established between word-final *-n* deletion and vowel reduction in loanwords with a word-final *-n* preceded by a stressed vowel and a mid vowel in unstressed position (3) (for a more comprehensive study including other types of interactions, see Pons-Moll, Torres-Tamarit & Martin-Diaconescu 2019a, b).

In what follows we describe the two processes of Eastern Catalan under analysis (see §§ 2.1 and 2.2) and show how they can interact within the same loanword.

2.1 Word-final *-n* deletion and vowel reduction in Eastern Catalan

Word-final *-n* deletion (ND) and vowel reduction (VR) are general processes in the native lexicon of Eastern Catalan. This can be seen in the alternations in (3), where a stem-final alveolar nasal [n] alternates with [Ø] in word-final position and after a stressed vowel; and in the alternations in (4), where the stressed low vowel [á] and the stressed mid-front vowels [é] and [ê] alternate with [ə] in unstressed

position and where the stressed mid-back vowels [ɔ] and [ó] alternate with [u] also in unstressed position. As shown in (4), this process of VR affects vowels placed in any position within the word, and also explains the realization of non-alternating *a* and *e* as [ə] and of non-alternating *o* as [u] in unstressed position: *esperança* [əspə́ransə], *home* [ɔ́mə], *hospítal* [uʃpítal], *carro* [káru], etc. (Throughout, acute marks indicate stress.)

(3) ND in inherited words

(Mascaró 1976; Bonet & Lloret 1998; Faust & Torres-Tamarit 2017)

<i>plans</i> [pláns]	~ <i>planíssim</i> [plənísim]	~ <i>pla</i> [pláŋ]
“flat.PL”	~ “flat.SUPERL”	~ “flat.SG”
<i>cosins</i> [kuzíns]	~ <i>cosinet</i> [kuzinét]	~ <i>cosí</i> [kuzíŋ]
“cousin.PL”	~ “cousin.DIM”	~ “cousin.SG”

(4) VR in inherited words

(Mascaró 1976; Bonet & Lloret 1998)

<i>casa</i> [kázə]	~ <i>caseta</i> [kəzétə]
“house”	~ “house.DIM”
<i>terra</i> [térə]	~ <i>terrestre</i> [təréstrə]
“earth”	~ “terrestrial”
<i>esquerrà</i> [əskərá]	~ <i>esquerranisme</i> [əskəranízmə]
“leftist”	~ “leftism”
<i>fera</i> [fērə]	~ <i>feroç</i> [fəros]
“beast”	~ “fierce”
<i>porta</i> [pórtə]	~ <i>portal</i> [purtál]
“door”	~ “hallway”
<i>poma</i> [pómə]	~ <i>pomera</i> [pumérə]
“apple”	~ “apple tree”
<i>bastó</i> [bəstó]	~ <i>bastonet</i> [bəstunét]
“stick”	~ “stick.DIM”

2.2 Underapplication of ND and VR in loanwords

ND and VR, though, tend to underapply in loanwords, as shown in (5) and (6), independently of when in the history of the Catalan language were introduced (cf. *cancan* vs. *Mégane*) and of the donor language (cf. *futon* from Japanese vs. *paston* from Spanish; *gadget* from English vs. *ramen* from Japanese). In (5) we find examples with an alveolar nasal in the context of application of ND: that is, in word-final position and after a stressed vowel. In (6) we find examples with [e] and with [o] in the context of application of VR: that is, in unstressed position.

(5) Underapplication of ND in loanwords

(Pons-Moll 2012, 2015; Pons-Moll et al. 2019a; b)

<i>divan</i>	[diβá <u>́</u> n]	“divan”
<i>futon</i>	[futó <u>́</u> n]	“futon”
<i>cancan</i>	[kaŋká <u>́</u> n]	“cancan”
<i>xaman</i>	[ʃomá <u>́</u> n]	“shaman”
<i>catamaran</i>	[kətəmərə <u>́</u> n]	“catamaran”
<i>taliban</i>	[təliβá <u>́</u> n]	“taliban”
<i>catipén</i>	[kətipé <u>́</u> n]	“odor, slang”
<i>maton</i>	[mətó <u>́</u> n]	“bully”
<i>paston</i>	[pəstó <u>́</u> n]	“a bundle”
<i>Pakistan</i>	[pəkistá <u>́</u> n]	“Pakistan”
<i>Afganistan</i>	[əvɣənistá <u>́</u> n]	“Afganistan”
<i>Sudan</i>	[suðá <u>́</u> n]	“Sudan”
<i>Vuitton</i>	[bujtó <u>́</u> n]	“Vuitton”
<i>Nissan</i>	[nisá <u>́</u> n]	“Nissan”

(6) Underapplication of VR in loanwords

(Mascaró 2002; Cabré 2009; Pons Moll 2012; Pons-Moll et al. 2019a; b)

<i>cutre</i>	[kútr <u>̄</u> e]	“crap”
<i>gore</i>	[góru <u>̄</u> e]	“gore”
<i>flyer</i>	[fláj <u>̄</u> er]	“flyer”
<i>gadget</i>	[gádʒ <u>̄</u> et]	“gadget”
<i>hí́ster</i>	[xíp <u>̄</u> ster]	“hipster”
<i>mà́ster</i>	[má <u>̄</u> ster]	“master”
<i>cú́ter</i>	[kú <u>̄</u> ter]	“cuter”
<i>blí́ster</i>	[blí <u>̄</u> ster]	“blister”
<i>Twitter</i>	[twít <u>̄</u> er]	“Twitter”
<i>youtuber</i>	[jutú <u>̄</u> ber]	“youtuber”
<i>tempura</i>	[təmpú <u>̄</u> rə]	“tempura”
<i>serotonina</i>	[serot <u>̄</u> onínə]	“serotonin”
<i>contà́iner</i>	[k <u>̄</u> ontájner]	“container”
<i>mojito</i>	[moxít <u>̄</u> o]	“mojito”
<i>pesto</i>	[pést <u>̄</u> o]	“pesto”
<i>judo</i>	[ʒú <u>̄</u> ðo]	“judo”
<i>sado</i>	[sá <u>̄</u> ðo]	“sado”
<i>crono</i>	[krón <u>̄</u> o]	“chrono”
<i>taco</i>	[tá <u>̄</u> ko]	“taco”
<i>pà́kinson</i>	[pá <u>̄</u> rkinson]	“Parkinson”
<i>gastrobar</i>	[gastro <u>̄</u> βár]	“gastrobar”
<i>collage</i>	[kolá <u>̄</u> ʃ]	“collage”

2.3 Interaction of ND and VR in the same loanword

Interestingly, according to the results of our production and judgment tests (see § 3), loans susceptible to undergoing both processes (7) show a consistent behavior in which underapplication of the two processes is the most common solution (Pat1: $t[o]b[o]ga[n]$; 8a), followed closely by just underapplication of ND (Pat2: $t[u]b[u]ga[n]$; 8b), followed a long way behind by application of both processes (Pat3: $t[u]b[u]ga[\emptyset]$; 8c), and in which underapplication of VR and application of ND (Pat4: $*t[o]b[o]ga[\emptyset]$; 8d) is unattested. From now on, we illustrate our arguments with the loan *tobogan* (which is, in fact, an old loanword), although the patterns described can be extended to all loans with a parallel structure (see 7). (Note that in the examples in 7, only the stressed vowel is transcribed, because, as noted, these forms can be realized in diverse ways; see 8.)

- (7) Convergence of word-final *-n* and unstressed mid-vowels in Catalan loanwords

tobog[á]n	“toboggan”
orangut[á]n	“orangutan”
wont[ó]n	“wonton”
OT[á]N	“OTAN”
xarlest[ó]n	“charleston”
sed[á]n	“Sedan”
Pequ[í]n	“Beijing”
Berl[í]n	“Berlin”
leviat[á]n	“leviathan”
fiest[ó]n	“big party”
Meg[á]ne	“car brand”
Teher[á]n	“Teheran”
Vall d’Hebr[ó]n	“place name”
leviat[á]n	“leviathan”
Decathl[ó]n	“sports shop”
Redox[ó]n	“medicine brand”

- (8) Interaction of ND and VR in loanwords

Most common (non-nativized)	Underapplication of ND and VR	$t[o]b[o]ga[n]$	Pat1
Less common (partially nativized)	Underapplication of ND and application of VR	$t[u]b[u]ga[n]$	Pat2
Least common (fully nativized)	Normal application of ND and VR	$t[u]b[u]ga[\emptyset]$	Pat3
Unattested (partially nativized)	Underapplication of VR and application of ND	$*t[o]b[o]ga[\emptyset]$	Pat4

That is to say, underapplication of both processes can co-occur, as can application of both processes and application of VR and underapplication of ND, but underapplication of VR and application of ND cannot. As for the implicational relations established between these processes, we can thus state that if ND applies so does VR, but not vice versa; in parallel, if VR is blocked so is ND, but not vice versa.

3. Experimental surveys

These generalizations are drawn from two experimental surveys, a production test (§ 3.1) and a perception test (§ 3.2), carried out with 31 Barcelona Catalan speakers aged 18–23 in 2017–2018. The speakers were students at the Universitat de Barcelona, most of them pursuing the BA degree in *Comunicació i Indústries Culturals*. The selected students were all born in Barcelona, with parents also from Barcelona, and with Catalan as their usual language. The students were recruited from the courses *Gèneres i Formats de la Comunicació Escrita*, *Llengua Catalana II*, and via Facebook advertisements.

3.1 Picture-naming production test

In the production task, speakers were asked to utter 16 loanwords, for which referent pictures were provided in a PowerPoint. All loanwords presented the relevant structures, that is, soliciting words with a word-final /n/ after a stressed vowel and the unstressed mid vowels *e* and *o* or the low vowel *a* (*tobogan*, *caiman*). The test was completed with loanwords with just one of the relevant structures, that is, loans either with a word-final nasal after a stressed vowel or with a mid vowel in unstressed position (i.e., *divan*, *màster*, respectively), and with 50% distractors, and items were randomized. The results of this test, which can be seen in (9), indicate that speakers produced these loanwords following Pat1 in around 65% of the cases, Pat2 in 25%, and Pat3 in around 10%. No speaker produced these loanwords following Pat4. The table in (9) reflects the results for loans with unstressed low vowel *a* and loans with the unstressed mid-vowels *e* or *o* altogether. Considered separately, it can be observed that vowel reduction is more frequent when the affected vowel is low (as in *caiman*) than when it is mid (as in *tobogan*), as shown in the table in (10); the differences, though, are less noticeable than expected, since vowel reduction of the low vowel has traditionally been considered compulsory – unlike vowel reduction of the mid vowels, which has more exceptions and behaves differently in loanwords. Note, in this respect, that the speakers showed regular vowel reduction of *a* to schwa in filler native words.

- (9) Results of the picture-naming production test (including loans with the combination of the unstressed mid vowels *e* and *o*, the low vowel *a* and a final *-n*: *tobogan*, *caiman*, etc.)

Patterns	% of answers
a. _{PAT1} t[o]b[o]ga[n], c[a]ima[n]	65.2%
b. _{PAT2} t[u]b[u]ga[n], c[ə]ima[n]	25%
c. _{PAT3} t[u]b[u]ga[Ø], c[ə]ima[Ø]	9.8%
d. _{PAT4} t[o]b[o]ga[Ø], c[a]ima[Ø]	0%

- (10) Results of the picture-naming production test (including only loans with the combination of the unstressed low vowel *a* and a final *-n*: *caiman*, etc.)

Patterns	% of answers
a. _{PAT1} c[a]ima[n]	54%
b. _{PAT2} c[ə]ima[n]	37.8%
c. _{PAT3} c[ə]ima[Ø]	8.2%
d. _{PAT4} c[a]ima[Ø]	0%

3.2 Judgment test

The same 31 speakers were asked to evaluate the naturalness of the four possible patterns for the same 16 loanwords (22×4 patterns = 88 items) on a Likert scale of 1–5 according to the following categories: very natural, fairly natural, neutral, fairly unnatural, very unnatural. The test was presented to the speakers in an audio file via a Google form on the Internet. The test was also completed with loanwords with just one of the relevant structures (i.e., *divan*, *màster*) and with 50% distractors, and randomized. The results of the test, shown in (11), are slightly more variable than those in the production test, but some consistent tendencies can be observed. Pat1 received a high score for the very natural and fairly natural categories, with Pat2 not far behind. Pat3 and Pat4, in contrast, received very low scores for these categories, and a high score for the very unnatural and fairly unnatural categories. Note that no significant differences were detected with respect to the quality of the unstressed vowels (i.e., low /a/, as in *c[a]iman*, vs. mid /e/, /o/, as in *t[o]b[o]gan*), as shown in the tables of (12).

- (11) Results of the judgment tests (including loans with the combination of the unstressed mid vowels *e* and *o*, the low vowel *a* and a final *-n*: *tobogan*, *caiman*, etc.)

<small>PAT1</small> t[o]b[o]ga[n], c[a]ima[n]	% of answers
Very natural	47.6%
Fairly natural	17.3%
Neutral	14.1%
Fairly unnatural	15.1%
Very unnatural	5.8%

<small>PAT2</small> t[u]b[u]ga[n], c[ə]ima[n]	% of answers
Very natural	35.5%
Fairly natural	25.6%
Neutral	17.7%
Fairly unnatural	15.7%
Very unnatural	5.4%

<small>PAT3</small> t[u]b[u]ga[Ø], c[ə]ima[Ø]	% of answers
Very natural	12.0%
Fairly natural	16.0%
Neutral	16.8%
Fairly unnatural	27.1%
Very unnatural	28.1%

<small>PAT4</small> t[o]b[o]ga[Ø], c[a]ima[Ø]	% of answers
Very natural	10.9%
Fairly natural	15.1%
Neutral	15.9%
Fairly unnatural	25.2%
Very unnatural	32.9%

- (12) Results of the judgment tests (distinguishing loans with the combination of the unstressed mid vowels *e* and *o* and a final *-n* and loans with the low vowel *a* and a final *-n*)

<small>PAT1</small> t[o]b[o]ga[n]	% of answers	<small>PAT1</small> c[a]ima[n]	% of answers
Very natural	42.9%	Very natural	51.3%
Fairly natural	13.8%	Fairly natural	20.1%
Neutral	15.2%	Neutral	13.3%
Fairly unnatural	20.3%	Fairly unnatural	11.1%
Very unnatural	7.8%	Very unnatural	4.3%

^{PAT2} t[u]b[u]ga[n]	% of answers
Very natural	26.3%
Fairly natural	22.1%
Neutral	21.2%
Fairly unnatural	23.5%
Very unnatural	6.9%

^{PAT2} c[ə]ima[n]	% of answers
Very natural	42.7%
Fairly natural	28.3%
Neutral	15.1%
Fairly unnatural	9.7%
Very unnatural	4.3%

^{PAT3} t[u]b[u]ga[Ø]	% of answers
Very natural	10.6%
Fairly natural	13.9%
Neutral	18.5%
Fairly unnatural	26.9%
Very unnatural	30.1%

^{PAT3} c[ə]ima[Ø]	% of answers
Very natural	13.3%
Fairly natural	17.6%
Neutral	15.4%
Fairly unnatural	27.2%
Very unnatural	26.5%

^{PAT4} t[o]b[o]ga[Ø]	% of answers
Very natural	7.8%
Fairly natural	8.8%
Neutral	14.3%
Fairly unnatural	26.7%
Very unnatural	42.4%

^{PAT4} c[a]ima[Ø]	% of answers
Very natural	13.3%
Fairly natural	20.1%
Neutral	17.2%
Fairly unnatural	24.0%
Very unnatural	25.4%

These results, which for the most part conform to the gradations presented in § 2, deserve a number of comments. First, again we attribute the low scores for Pat3 *t[u]b[u]ga[Ø]* (i.e., nativized patterns) to the age of the speakers interviewed: it is well-known that older speakers are more conservative (e.g., Labov 1980) and tend to make utterances that conform to the native phonology (fully nativized patterns), whereas younger speakers tend to follow the non-nativized patterns. Second, as already clarified, no significant differences were detected with respect to the quality of the unstressed vowels, especially in the judgment test. Third, as expected, the results are more conclusive in the production test than the judgment test, where there is more variability: whereas the production test reproduces the actual utterances of speakers, the judgment test reproduces the utterances of the speakers themselves but also the ones they are used to hearing in their speech community, that is, the utterances they are familiar with.

We conclude that the results of the judgment test reproduce *grosso modo* the production test results, except for the close relative well-formedness of Pat4 in

relation to Pat3. The reasons why we discard Pat4 and include Pat3 as possible realizations (see again 8) are that Pat4 received 0% of answers in the production test, whereas Pat3 received a 9.8%, and that the low scores for Pat3, both in the production and the judgment test, must be relativized, as noted, by the age of the speakers.


4. Analysis with weighted scalar constraints

Implicational patterns of the sort exemplified in the previous sections are expected to exist in a model with weighted constraints like Harmonic Grammar (see, among others, Smolensky & Legendre 2006; Pater 2009), and more specifically with Weighted Scalar Constraints, as developed in Hsu & Jesney (2017, 2018). In § 4.1 we briefly introduce Harmonic Grammar, and in § 4.2 we explain how Weighted Scalar Constraints can be applied to strata, in order to account for the possible and impossible nativizations under study and for the implicational relationships between processes depicted in the previous sections.

4.1 Harmonic grammar: A brief overview

According to Harmonic Grammar (Smolensky & Legendre 2006; Pater 2009), cross-linguistic variation is not explained by different constraint rankings (as in Optimality Theory), but by constraints with different weights. The violation of a constraint implies the assignment of a negative value, and this value is multiplied by the constraint weight: if a constraint has a weight of 5.5, its violation by a candidate implies the assignment of the negative value -5.5 ; if the candidate violates this constraint twice, the assignment will be -11 , and so on. The sum of the negative values obtained depending on the violations of the different constraints constitutes the harmony of a candidate. The winning candidate is the one that obtains the highest negative value, i.e., the lowest penalty. In (13), we illustrate HG with an example of vowel reduction in Catalan. The two competing constraints are $*e, o_{\text{UNSTR}}$, against mid vowels in unstressed position and with a weight of 5.5, and $\text{IDENT-}V_{\text{UNSTR}}$, protecting the underlying featural specification of vowels in unstressed position and with a weight of 2. The former outweighs the latter, and this explains why the candidate with vowel reduction (with a penalization of -2) is selected as optimal rather than the faithful candidate (with a more severe penalization of -5.5).

(13) HG illustrated through VR

$/\text{mez}+\epsilon\text{t}/$ month.DIM	$*e, o_{\sigma_{\text{UNSTR}}}$ $w = 5.5$	IDENT- V_{UNSTR} $w = 2$	H
a. [mezét]	-1		-5.5 (-1 × 5.5)
 b. [mezét]		-1	-2 (-1 × 2)

4.2 Weighted scalar constraints applied to strata

Weighted Scalar Constraints pivot on the ‘scaling factor’. Within HG, what the scaling factor does is to increase progressively and proportionately the weight of either faithfulness or markedness constraints. Thus, the weight relations between markedness and faithfulness constraints vary as the scaling factor for faithfulness or markedness increases and, at a certain scaling factor, these relations can vary so much that this results in the selection of a different candidate than the one selected with an inferior scaling factor. Applied to strata, the scaling factor determines the transition from one stratum to another and thus the different levels of nativization of the speaker’s lexicon (Hsu & Jesney 2017, 2018). The specific proposal of Hsu & Jesney (2018: 4) works as follows: “the penalty associated with the violation of a constraint is scaled based on distance from the core lexicon. The total penalty for a constraint violation is $w \times s(d)$, where w is the base constraint weight, s is the scaling factor, and d is a measure of distance from the core. Values for d begin at 0, in the case of words within the core lexicon, and increase as the degree of nativization decreases”. In order to account for the behavior of loanwords and for the organization of the lexicon into strata, it is faithfulness violations that are scaled, so their weight values increase from the nuclear stratum to the peripheral one. A general definition for the scaled version of faithfulness constraints is given in (14).

(14) Scaled Faithfulness

Given a basic constraint weight w ,
 a scaling factor s , and a distance from the core d ,
 for each input structure that is not realized faithfully in the output,
 assign a weighted violation score of $w \times s(d)$.

The virtue of this proposal, in contrast to Itô & Mester’s (see § 5), is that, given certain weights and given any scaling factor, there are certain weight constraint relations, and thus certain patterns, that are systematically and permanently excluded from the grammar. The model, thus, straightforwardly captures the asymmetrical relationships between processes observed in natural languages. This approach, for instance, is pursued in Hsu & Jesney (2017) to account for the possible patterns of

adaptation of loanwords of English in Quebec French. In Quebec French, English loans with a rhotic [ɹ] and a final affricate [tʃ] can be adapted by repairing both segments (*scratch* [skɾaʃ]), by just repairing the rhotic ([skɾatʃ]), and by not repairing any segment ([skɾaʃ]), but it is not possible to adapt them by repairing the affricate and not the rhotic (*[skɾaʃ]). That is, if the affricate is repaired, so is the rhotic, but this does not apply the other way around (Hsu & Jesney 2017: 250).

In what follows we present the modelling of the nativization patterns presented in § 2 and § 3 according to this framework.

4.3 Analysis of the interaction between ND and VR

For Catalan, we assume a structure made up of three lexical strata: (i) the core one, for speakers and loans with application of VR and ND ($t[u]b[u]ga[Ø]$); (ii) the intermediate one, for speakers and loans with just application of VR ($t[u]b[u]ga[n]$); (iii) the peripheral one, for speakers and loans with underapplication of both VR and ND ($t[o]b[o]ga[n]$). For all loanwords in all strata, we assume an underlying form with a low or a mid vowel and a word-final alveolar nasal (/kajman/; /tobogan/); this is because we understand that most of these loanwords are introduced into Catalan either through the written form (with a spelling with *e*, *o* and *a*, and a final *n*: *orangutan*) or through Spanish (where the structures involved are pronounced faithfully: [orangután]). The two markedness constraints involved are * e, o_{UNSTR} (against unstressed high-mid vowels) (15a) and * $n]_{\text{WD}}$ (against word-final posttonic -*n*) (15b), which receive stable weights of 5.5 and 2.5 respectively across all three possible strata. These two markedness constraints interact with the faithfulness constraints IDENT- V_{UNSTR} (against featural changes for unstressed vowels) (15c) and MAX-IO (against deletion) (15d), which receive stable weights of 2 and 1.5 respectively across all three possible strata. (Note that these weights have been calculated considering both the interaction of nasal deletion and vowel reduction and the interaction of vowel reduction and vowel laxing in Catalan loanwords; see Pons-Moll et al. 2019a; b, for evidence in this direction.)

- (15) a. * $n]_{\text{WD}}$: Assign one violation mark for every nasal in word-final position and after a stressed vowel.
 b. * e, o_{UNSTR} : Assign one violation mark for every unstressed high-mid vowel.
 c. MAX-IO: Assign one violation mark for every segment in the input that has no correspondent in the output.
 d. IDENT- V_{UNSTR} : Assign one violation mark for every unstressed vowel in the output whose input correspondent has a different featural specification.

As illustrated in the tableaux in (16), scaled faithfulness ensures that the weight values for the faithfulness constraints increase from the core stratum (in which $s = 1$), towards the intermediate stratum (which starts with $s = 1.8$), until reaching the peripheral stratum (which starts with $s = 2.8$). Faithfulness values thus acquire higher relevance the closer they are to the peripheral stratum. Given the constraint weights, no scaling factor can yield the impossible nativization $\text{Pat4} *t[o]b[o]ga[\emptyset]$ (as shown by the strata cross overpoints in 17, discussed below). In the ‘core stratum’, markedness constraints outweigh faithfulness constraints, which explains why the selected candidate is the one with the application of all ‘native’ processes. In the ‘intermediate stratum’, the scaling factor of 1.8 is enough for the constraint MAX-IO to outweigh the markedness constraint $*n]_{\text{WD}}$, with which it interacts, but not for the constraint $\text{IDENT-V}_{\text{UNSTR}}$ to outweigh $*e, o_{\text{GUNSTR}}$, and this explains the selection of the candidate with the mixed pattern (with vowel reduction but no word-final $-n$ deletion). In the ‘peripheral stratum’, the scaling factor of 2.8 is high enough for both faithfulness constraints to outweigh the markedness constraints with which they are in conflict, and this explains the selection of the pattern with underapplication of both processes and non-nativization. Note that, in order to maintain these predictions in which Pat4 is permanently excluded, the scaling factor cannot be applied to a single faithfulness constraint, or cannot differ depending on the faithfulness constraint involved: the scaling factor must affect MAX-IO and $\text{IDENT-V}_{\text{UNSTR}}$ as a whole, since the weight proportion between them must be maintained across strata.

At this point, it is important to note that the same speaker may associate one set of words with one stratum, another set with another stratum, and yet another set with a third stratum. That is, there are two factors that determine the selection of patterns: the speaker’s grammar itself and the stratum to which each set of words is associated in the grammar of this speaker. The tableau in (16), thus, abstractly reproduces the nativization patterns available in Catalan grammar.

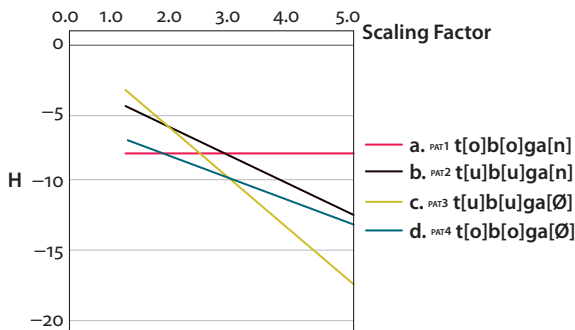
Summarizing what we have said so far, in our proposal all markedness constraints are satisfied in the core stratum, only the markedness constraint $*e, o_{\text{GUNSTR}}$ in the intermediate stratum, and finally none of the markedness constraints in the peripheral stratum.

(16) HG with ‘weighted scalar constraints’ tableau for the interaction between VR and ND

<i>i.</i> /tobogan/	*e, o _{σ_{UNSTR}} w = 5.5	*n] _{WD} w = 2.5	IDENT-V _{UNSTR} w = 2	MAX-IO w = 1.5	H	Scaling factor for F	Strata
a. [toβoγán]	-2	-1			-13.5 (-11)+(-2.5)	1	Core stratum
b. [tuβuyán]		-1	-2		-6.5 (-2.5)+(-4)		
☞ c. [tuβuyáØ]			-2	-1	-5.5 (-4)+(-1.5)		
d. [toβoγáØ]	-2			-1	-12.5 (-11)+(-1.5)		
<i>ii.</i> /tobogan/	*e, o _{σ_{UNSTR}} w = 5.5	*n] _{WD} w = 2.5	IDENT-V _{UNSTR} w = 2	MAX-IO w = 1.5	H	Scaling factor for F	
a. [toβoγán]	-2	-1			-13.5 (-11)+(-2.5)	1.8	Intermediate stratum
☞ b. [tuβuyán]		-1	-2		-9.7 (-2.5)+(-4 × 1.8)		
c. [tuβuyáØ]			-2	-1	-9.9 (-4 × 1.8)+(-1.5 × 1.8)		
d. [toβoγáØ]	-2			-1	-13.7 (-11)+(-1.5 × 1.8)		
<i>iii.</i> /tobogan/	*e, o _{σ_{UNSTR}} w = 5.5	*n] _{WD} w = 2.5	IDENT-V _{UNSTR} w = 2	MAX-IO w = 1.5	H	Scaling factor for F	
☞ a. [toβoγán]	-2	-1			-13.5 (-11)+(-2.5)	2.8	Peripheral stratum
b. [tuβuyán]		-1	-2		-13.7 (-2.5)+(-4 × 2.8)		
c. [tuβuyáØ]			-2	-1	-15.4 (-4 × 2.8)+(-1.5 × 2.8)		
d. [toβoγáØ]	-2			-1	-15.2 (-11)+(-1.5 × 2.8)		

In (17), finally, we present the chances of selection of each pattern given the weights assigned to each constraint and depending on the scaling factor. The most important point here is that Pat4 (the blue line) is always harmonically bounded by some of the other patterns, so there is no chance for this pattern to be selected in the language. Pat1, the pink line, harmonically bounds Pat2 and Pat3 at any scaling factor equal or higher than 2.8; Pat2, the black line, does so at the scaling factor interval {1.8–2.7} and Pat3, the ocre line, at the scaling factor interval {1–1.7}.

(17) Strata cross overpoints



5. Alternative analyses

The type of implicational patterns analyzed in this paper have been accounted for in previous work by resorting either to the indexation of constraints that apply to individual lexical strata (Itô & Mester 1995, 1999, 2008) or to separate co-phonologies associated with individual lexical strata (Inkelas & Zoll 2007). These approaches predict all possible patterns, but nothing prevents the ‘overgeneration’ of the impossible ones: given intrinsic OT constraint reranking (in this specific case across strata or across co-phonologies), nothing prevents rankings such as, for instance, $*n]_{WD} \gg \text{MAX-IO}, \text{IDENT-}V_{UNST} \gg \text{IDENT-}V_{UNST}$, from leading to the impossible pattern $*t[o]b[o]ga[\emptyset]$. This is why Itô & Mester (1999) resort to the metacondition “Ranking consistency”, which “forces” certain ranking relations between markedness and faithfulness, and which therefore ensures the same rankings across strata: “Let F and G be two types of I-O faithfulness constraints ..., there are no strata A, B such that the relative rankings of the indexed versions of F and G are inconsistent with each other. If $F/A \gg G/A$ for some stratum A, then there is no stratum B such that $G/B \gg F/B$ ” (p. 27). According to Itô & Mester (1999: 28), “[t]here is an underlying unity behind the various stratal incarnations of a given faithfulness constraint”. As shown in this paper, metaconditions are not necessary within Harmonic Grammar

with Scalar Constraints, where the weight of the constraints, along with any scaling factor, gives no chance to the impossible patterns (*t[o]b[o]ga[Ø], *c[a]ima[Ø]).

6. Concluding remarks

In this paper we have explored phonological nativization patterns in Catalan loanwords, and have shown, on the basis of production and judgment tests, that the two processes under scrutiny (word-final *-n* deletion and vowel reduction) interact in an asymmetrical way. We have argued that these asymmetrical interactions can be straightforwardly formalized by using Harmonic Grammar with Scalar Weighted Constraints (Hsu & Jesney 2017, 2018), in which faithfulness constraints acquire an increasing relevance from the core stratum to the peripheral stratum and in which, if a process fails to apply in a given stratum, it will also fail to apply in more peripheral strata, but not the other way around. Future research should seek to verify whether similar predictions are borne out by descaling the weight of markedness constraints from the nuclear stratum to the peripheral one, or by scaling the weight of the markedness constraints from the peripheral stratum to the nuclear one. It should also be investigated whether considering the whole lexicon of Catalan (including, therefore, native and non-native words), the preferred factor is the closest to 1 and whether there is indeed an attraction towards this factor in the process of nativization of the non-native lexicon. Finally, from a sociolinguistic point of view, it is interesting that young speakers in the Barcelona area show a clear preference for patterns with a lack of nativization and thus far from regular phonology, which we interpret as a consequence of the innovative nature of the speech of this group of speakers (Labov 1980; Milroy & Milroy 1985). In future studies, it will be necessary to explore the extent to which the same trends are detected in speakers of other areas of the Catalan-speaking territory and pertaining to older age groups.

Acknowledgements

We thank two anonymous reviewers and the editors for their constructive comments, Vlad Martin-Diaconescu, who did the Excel calculations of the weights depending on the scaling factor variable, and the audiences of Going Romance 2018 (December 2018), GLOW 42 (May 2019), and the 27 Manchester Phonology Meeting (May 2019).

Funding

This paper has been supported by the projects FFI2016-76245-C3-3-P (Spanish Ministry of Economy and Competitiveness and European Regional Development Fund) and 2014SGR918 (Catalan Government), and it is done as part of the research group GEVAD (<http://www.ub.edu/GEVAD/>).

References

- Bonet, Eulàlia & Maria-Rosa Lloret. 1998. *Fonologia catalana*. Barcelona: Ariel.
- Cabré, Teresa. 2009. Vowel reduction and vowel harmony in Eastern Catalan loanword phonology. In Marina Vigário, Sónia Frota & Maria Joao Freitas (eds.), *Phonetics and phonology. Interactions and interrelations*, 267–286. Amsterdam: John Benjamins.
<https://doi.org/10.1075/cilt.306.13cab>
- Faust, Noam & Francesc Torres-Tamarit. 2017. Stress and final /n/ deletion in Catalan: Combining Strict CV and OT. *Glossa: a Journal of General Linguistics* 2(1). 63. 1–23.
<https://doi.org/10.5334/gjgl.64>
- Hsu, Brian & Karen Jesney. 2017. Loanword adaptation in Québec French: Evidence for weighted scalar constraints. In Aaron Kaplan, Abby Kaplan, Miranda K. McCarvel & Edward J. Rubin (eds.), *Proceedings of the 34th Meeting of the West Coast Conference on Formal Linguistics (WCCFL 34)*, 249–258. Somerville, MA: Cascadilla.
- Hsu, Brian & Karen Jesney. 2018. Weighted scalar constraints capture the typology of loanword adaptation. In Gillian Gallagher, Maria Gouskova & Sora Heng Yin (eds.), *Proceedings of the 2017 Annual Meeting on Phonology*, 1–11. Linguistic Society of America.
<https://doi.org/10.3765/amp.v5io.4246>
- Inkelas, Sharon & Cheryl Zoll. 2007. Is grammar dependence real? A comparison between cophological and indexed constraint approaches to morphologically conditioned phonology. *Linguistics* 45. 133–171. <https://doi.org/10.1515/LING.2007.004>
- Itô, Junko & Armin Mester. 1995. The core-periphery structure of the lexicon and constraints on reranking. In Jill Beckman, Suzanne Urbanczyk & Laura Walsh (eds.), *University of Massachusetts Occasional Papers in Linguistics [UMOP] Volume 18: Papers in Optimality Theory*. 181–209. Amherst, MA: GLSA.
- Itô, Junko & Armin Mester. 1999. The phonological lexicon. In Natsuko Tsujimura (ed.), *The handbook of Japanese linguistics*, 62–100. Malden, MA: Blackwell. [Also in John J. McCarthy (ed.) (2004). *Optimality Theory: A reader*, 552–568. Malden, MA: Wiley Blackwell.]
- Itô, Junko & Armin Mester. 2008. Lexical classes in phonology. In Shigeru Miyagama & Mamoru Saito (eds.), *The Oxford handbook of Japanese linguistics*, 84–106. Oxford: Oxford University Press.
- Labov, William. 1980. *Locating language in time and space*. New York: Academic Press.
- Mascaró, Joan. 1976. Catalan phonology and the phonological cycle. Ph.D. Dissertation: MIT. Published by Indiana University Linguistics Club, Bloomington, 1978.
- Mascaró, Joan. 2002. El sistema vocàlic. Reducció vocàlica. In Joan Solà, Maria-Rosa Lloret, Joan Mascaró & Manuel Pérez Saldanya (eds.), *Gramàtica del català contemporani*, 89–123. Barcelona: Empúries.

- Milroy, James & Lesley Milroy. 1985. Linguistic change, social network and speaker innovation. *Journal of Linguistics* 21. 339–384. <https://doi.org/10.1017/S0022226700010306>
- Pater, Joe. 2009. Weighted Constraints in Generative Linguistics. *Cognitive Science* 33. 999–1035. <https://doi.org/10.1111/j.1551-6709.2009.01047.x>
- Pons-Moll, Clàudia. 2012. Loanword phonology, lexical exceptions, morphologically driven underapplication, and the nature of positionally biased constraints. *Catalan Journal of Linguistics* 11. 127–166. <https://doi.org/10.5565/rev/catj1.13>
- Pons-Moll, Clàudia. 2015. Comentaris a “Regularitat i excepcions en fonologia: les reduccions vocàliques”, de Joan Mascaró. In Maria-Rosa Lloret, Clàudia Pons-Moll & Eva Bosch-Roura (eds.), *Clàssics d'ahir i d'avui en la gramàtica del català*, 71–101. Barcelona: Publicacions i Edicions de la Universitat de Barcelona.
- Pons-Moll, Clàudia, Francesc Torres-Tamarit & Vlad Martin-Diaconescu. 2019a. Catalan nativization patterns in the light of Weighted Scalar Constraints. Paper presented at Glow 42, Oslo, Norway, May 7–11.
- Pons-Moll, Clàudia, Francesc Torres-Tamarit & Vlad Martin-Diaconescu. 2019b. Catalan (im)possible nativizations in the light of Weighted Scalar Constraints. Paper presented at the 27th Manchester Phonology Meeting, Manchester, UK, May 23–25.
- Smolensky, Paul & Géraldine Legendre. 2006. *The harmonic mind: From neural computation to Optimality-Theoretic Grammar*. Cambridge, MA: MIT Press.

Temporal marking and (in)accessibility in Capeverdean

Fernanda Pratas
University of Lisbon

Recent descriptions have argued that what seem to be past tense markers in Capeverdean, a Portuguese-related language spoken in Cabo Verde, are instead allomorphs of a temporal agreement morpheme (Pratas 2018a). The rationale for this goes as follows. It is true that both *-ba*, from the variety of Santiago, and the related (and more complex) form *tava*, from São Vicente, are sometimes associated with a past tense in the terms of Klein (1994): the Topic Time is located before the Time of Utterance (Pratas 2014). This is the case in (i) past progressives and past habituais. But they also appear in (ii) subordinate clauses where no past interpretation is certain, such as some conditionals and other modal contexts. Since this subordinate lexical item is often licensed in the context of past situations denoted by their respective main clauses, it seems indeed better accounted for by this recent agreement proposal. That analysis, however, still leaves open the question of what this morpheme agrees with, and this is even more intriguing when it occurs fully separated from past situations. Alternatively, the approach taken in Pratas (2021) identifies a common point between (i) and (ii): all these structures denote situations with a low degree of accessibility from the speaker's perspective. This (in)accessibility is perceived in terms of time: in the first case, we cannot go back to the past; in the second case, external factors may (have) provide(d) an (in)accessible time location. The main goal of this paper then is to further defend this novel insight on that apparent mismatch, which can bring clues to similar problems crosslinguistically.

Keywords: temporal meaning, mood, aspect, accessibility, micro-variation

1. Introduction

This paper deals with the puzzling distribution of two variants of a temporal morpheme in Capeverdean, the Portuguese-related language spoken in Cabo Verde. One of these allomorphs, the postverbal *-ba* in the Santiago island variety, has been previously identified as an anterior tense marker (Suzuki, manuscript; Baptista

2002) and as a past marker (Pratas 2007, 2012, 2014). It often occurs, however, in embedded clauses where it has no past meaning, such as some conditionals and some purpose adverbial clauses. Similarly, the preverbal and more complex *tava*, which exists in São Vicente and brings together *ta* and *va* (see § 3.2.2 for an explanation), often marks past progressives and past habituais and also appears regularly in nonpast structures, such as conditionals and imagined situations.

This common connection of those forms with some nonpast meanings is amongst the arguments presented in this paper regarding its two main goals: (i) to summarize the proposal that what seem to be past markers in the language are in fact temporal agreement morphemes (Pratas 2018a); and (ii) to provide new evidence from Portuguese that confirm the crosslinguistic potential of the idea advanced in Pratas (2021) that many constructions traditionally taken as being morphologically marked for past are rather associated with a low degree of accessibility; in other words, the clauses in which they participate denote situations that are considered inaccessible from the speaker's point of view (some past situations and conditionals, among others), as opposed to those viewed as accessible (present and future situations). And so, in Capeverdean, this notion immediately puts together both past habituais and past progressives (no accessibility whatsoever), on the one hand, and some conditionals and other modal contexts (several levels of low accessibility), on the other hand. In the former, the inaccessibility corresponds to the past meaning of the whole sentence, where the said morpheme is also involved. In the latter, the temporal meaning is not necessarily past, but the same morpheme is there, equally expressing the low accessibility of the situation in question.¹

Note that the notion of accessibility used in this novel approach is different from the one in Arregui (2005), which determines a relation between tense and aspect in English *would*-conditionals and the nature of the worlds accessible (available) to the modal for quantification. The distinct use of this notion here is twofold: (i) here the accessibility relation is established not between possible worlds and a modal, but rather between the situation denoted by the sentence and the speaker; this is a deictic relation in nature, defined by the speaker according to their views about time: it is a well-known fact under different philosophical theories of time that the past is perceived by the human mind as no longer accessible (although we

1. This is different than what has been summarized in Binnick (2010): “some linguists have proposed defining the non-present tenses not in terms of temporal precedence or sequence but rather of detachment, understood as either detachment from the present or detachment from reality” (2010: 515). This is also different than what is been proposed in Klein (2010) as “non-canonical usages” of the present (2010: 48–51). The distinctive point here is that these morphemes are not viewed as tense morphemes, but rather direct markers of low accessibility – and so their uses are indeed canonical.

can think of it and ‘re-live’ it in our memories, we cannot return to it – time-travel is still viewed as impossible), whereas the present and the future are accessible (we are living it in the first case, and are headed for it in the second);² (ii) it extends to domains not traditionally considered as modals, like progressives in the past.

This common treatment of constructions which have been treated under a tense and aspect analysis and others which are easily taken as having a modal meaning represents a novel approach to these intriguing Capeverdean data. And if this analysis is on the right track, it will be worth testing also for other temporal morphemes in other languages.

The structure of the paper is as follows. Section 2 describes the main methodological and theoretical assumptions underlying the following discussion. Section 3 summarizes the proposal advanced in Pratas (2018a) on temporal morphemes in the main varieties under discussion here, spoken in the islands of Santiago and São Vicente. That proposal states that the Santiago postverbal morpheme *-ba* is not an anterior or a past tense marker, but rather a temporal agreement morpheme, and the same holds for the corresponding variant in São Vicente. This section is enriched with evidence from modal contexts. Finally, § 4 presents more arguments in favour of my current line of inquiry which takes these allomorphs to mark instead some degree of temporal inaccessibility: all these clauses denote situations which the speaker sees as more or less inaccessible; this latter analysis has also been defended in Pratas (2021) and is here improved with the inclusion of evidence from Portuguese. Section 5 concludes the paper.

2. Methodological and theoretical assumptions

This section describes Capeverdean in terms relevant for my current purposes (§ 2.1), and then presents some theoretical (§ 2.2) and methodological assumptions (§ 2.3) regarding linguistic variation.

2.1 Brief sociohistorical notes

Capeverdean emerged in the former Portuguese colony of Cabo Verde and inherited most of its lexicon from Portuguese, also revealing influence from West African languages, especially Mandinka and Wolof. It has been conventionally acknowledged as a creole, but since this classification has been the source of persistent

2. See Pratas (2021) for details regarding this relation between human conceptions of time and the linguistic expression of these values.

misconceptions (see Mufwene 2010 and Aboh & DeGraff 2017), it is here simply called Capeverdean.

The language now has two main varieties, roughly associated with the two most populated islands of the archipelago: (i) the Sotavento variety, which developed in the 16th century, in Santiago, and then in two other southern islands, Fogo and Brava (Lang 2014 and references therein), from contact between Portuguese and several different languages from the Mande and Atlantic families spoken by African slaves brought there by Portuguese slave traders; and (ii) the Barlavento variety, which developed in the 17th century in Santo Antão, the biggest and most productive among the northern islands, mainly peopled by migrants from the south (Swolkien & Cobbinah 2019 and references therein), and was further influenced by contact with Portuguese, due to the continuously arriving Portuguese settlers, among migrants from various other European countries. São Vicente was the last island to be populated, which only succeeded in the early 19th century, mainly (but not only) by migrants from other northern islands (Swolkien 2015 and references therein). These varieties are thus also spoken in other islands, with some variation of their own, but the main data presented here are from these two.

Capeverdean is not an official language in its own country, despite intense political debate. There is no standard variety, and attempts to establish an official spelling are still delayed by several problems. This absence produces interesting results when speakers write in Capeverdean, such as in social networks and blog posts, which deserve their own dedicated study (Gillier 2019).

2.2 Three main theoretical assumptions

The case of linguistic variation under analysis here is approached under three scientific assumptions. A general assumption is that language variation involves “alternative ways of ‘saying the same thing’” (Labov 1969: 738 fn. 20). There are then two other specific assumptions, which identify the theoretical framework used here, the generative enterprise in some of its more contemporary approaches: (i) the Borer-Chomsky conjecture, as it is called and stated in Baker (2008: 156): “All parameters of variation are attributable to differences in features of particular items in the lexicon (e.g., the functional heads)”; (ii) the notion that variation, in the ‘Labovian’ sense, involves “underspecification in the mapping between [functional] categories and morphological forms” (Adger & Smith 2010: 1109). This means that the same semantic features (of a given functional head in the lexicon) may be spelled out by different morphological units.

2.3 Methodology

The data analysed here were gathered during my own fieldwork, either in recorded semi-informal interviews or in programmed elicitation sessions, both in Santiago, with several fieldtrips between 2001 and 2016, and in São Vicente, during my first trip to this island at the end of January 2018. The mixing of examples from these semi-spontaneous conversations and experimental data – that is, specifically elicited for this research – is inspired by Matthewson (2004), on semantic fieldwork: although “elicitation ... does not involve direct inquiry about meaning”, we still need to confirm, the best we can, the exact meaning of the speakers’ utterances. The “fieldworker must obtain indirect clues to truth conditions and to felicity conditions” (Matthewson 2004: 379–380). I use two of these indirect means: (i) creating a scenario/context and then asking for the translation of the Portuguese sentence which is felicitous in that scenario; (ii) creating a scenario/context and asking for an acceptable linguistic production in that scenario.

3. Interactions of tense, aspect and mood

In this section, I summarize some key points about the crucial role played by aspect in this language (§ 3.1), provide empirical evidence as to why none of the available morphemes is a tense marker (§ 3.2), and describe my recent integrated proposal for all the combinations (§ 3.3).

3.1 The perfect and the progressive: The non-modal domain

The examples in (1a) and (1b) illustrate an interesting puzzle which has been the topic of several studies and subject of much debate (see Silva 1985, 1990; Suzuki manuscript; Veiga 1995; Baptista 2002; Swolkien 2015, among others).

- (1) a. *N sata/tita bebi/bebê agu.*
 1SG SATA/TITA drink water
 “I am drinking water.”
- b. *N bebi/bebê agu.*
 1SG drink water
 Intended meaning: “I drank (the) water.”

The past interpretation of the bare dynamic predicate in (1b) would be very difficult to explain through any other means than a perfect-like reading, in which case the

meaning is:³ “You have drunk (the) water.” Namely, it cannot be argued that the bare verb is inherently past, since if it were the case, we would have a past progressive in (1a) – the same verb form marked by the preverbal progressive morpheme. But it turns out that progressives in the past are different, as we see in (2).

- (2) (Santiago)
- a. [...] *Bu sata bebeba agu.*⁴
 2SG SATA drink:BA water
 (São Vicente)
- b. [...] *Bo tava ta bebê agu.*
 2SG TAVA TA drink water
 “You were drinking water.”

The perfect-like interpretation in (1b) is indeed in complementary distribution with the progressive meaning in (1a), which maps onto slightly different morphological forms: the preverbal *sata* or *aita* in Santiago, and *tita* or *tite* in São Vicente.⁵ This differs from the observation in typological studies for which “perfects and progressives are overwhelmingly periphrastic” (Dahl & Velupillai 2005: 2). In most cases in English, for instance, both progressive and perfect values are indeed obtained through an auxiliary + main verb (*be+V-ing* for the progressive, *have+participle* for the perfect).⁶ In Capeverdean there are also a few periphrastic forms expressing various temporal meanings which involve auxiliaries, but these are mostly used when a further value is added. In Santiago, this is the case of the habitual progressive: *ta+sta+ta+V* (in contexts like *N ta sta (senpri) ta durmi oki bu ta txiga* “I’m (always) sleeping when you arrive”). In São Vicente, there is also a periphrastic form for the progressive that combines with similar layers of aspectual / modal meanings (*ta+stod+ta+V*).

As semantically complex categories, the progressive and the perfect involve certain temporal characteristics (Smith 1991). Moreover, unlike the English Perfect, the perfect-like reading in Capeverdean is not incompatible with adverbials like *yesterday*. Still summarizing this analysis in Pratas (2018a), this perfect proposal

3. The event types assumed are the ones described in Vendler (1957). Activities, accomplishments and achievements are subsumed under the label ‘dynamic predicates’, as opposed to statives, which are not dynamic.

4. [...] is a notation for the obligatory context, which was not included here for reasons of space.

5. As in the proposals on linguistic variation in line with Adger & Smith (2010), as summarized in § 2.2.

6. See also Bertinetto, Ebert & de Groot (2000) for an overview of the progressive in European languages.

considers two different states resulting from the past situation, which still hold at the Topic Time (as defined in Klein 1994, see below):

- i. a resultant state, which is “an abstract state of the event’s ‘having occurred’” (Portner 2011: 1230) and is equivalent to the “post-time of the situation described by the predicates” (Klein 2014: 962). This is what we get with all dynamic predicates and with some types of stative predicates (therefore, for *kridita* ‘believe’ or *gosta* ‘like’, which have some stative and some non-stative interpretations, the resultant state is compatible with either case). According to the predicate in question, the resultant state/post-time of the situations described by the predicates has different implications regarding the current state of affairs, which does not interfere with the temporal reading of the situation;
- ii. a result state, which is part of the event structure (Moens & Steedman 1988; Smith 1991). This is what we get with some other statives, like *sta duenti* ‘be sick’ (stage-level), *e altu* ‘be tall’ (individual-level), some instances of modals like *podu* ‘can’, and some instances of *sabu* ‘know’, whose bare forms have a present interpretation. In this case, we have a past resultative situation of the type ‘get sick’, ‘get tall’, ‘get permission’, ‘get to know’, the temporal reading being directly anchored on the current result state (for a previous analysis of states in the language, see Pratas 2012, 2014).

The temporal interpretation through these aspectual distinctions is explained with the help of the following specifications. Three different Times are relevant to the construction of temporal reference, which were defined in Klein (1994) as: (i) Time of Utterance (TU); (ii) Situation Time (T-Sit); (iii) Topic Time (TT).⁷ This mediating time is defined by Klein as the time span about which an assertion is made (see also Stowell 2014 for discussion of this terminology, and cf. Demirdache & Uribe-Etxebarria 2000, et seq.). Klein later raises important questions about the notion of Time of Utterance (e.g., the speech event itself takes time; so, what slice of it are we considering?), but here it simply means ‘now’, the deictic temporal anchor of root clauses. The ordering of TU regarding TT thus corresponds to a past, a present or a future time reference, and the ordering of T-Sit regarding TT corresponds to a prospective aspect, the progressive, and the perfect, respectively.

7. Reichenbach (1947) proposed three time points – (i) Speech Time (S); (ii) Event Time (E); (iii) Reference Time (R) – and did not account for a relation later established as fundamental, which is inclusion. Also, the reference time/point was never well defined by the author. The terminology in Klein (1994) resolves these aspects and also includes a new distinction, i.e., between past perfective and past imperfective. This is not as salient here as the one between the perfect and the progressive, but the perfect has perfective features and the progressive is one among the imperfective meanings. For all these reasons, I adopt Klein’s terminology.

Therefore, the temporal orderings for Capeverdean dynamic predicates in root clauses are the following. Regarding aspect: (a) the bare form of the predicate means that T-Sit precedes TT – we get a perfect-like reading – and (b) a progressive morpheme means that T-Sit includes/coincides with TT. Regarding time reference: (c) when something in the context (note that past meanings always need a context) orders TU after TT, we have past perfect and past progressive readings, respectively, and (d) in the absence of that information, the coincidence between TU and TT is assumed through a pragmatic inference, thus corresponding to a present perfect and a present progressive. These relations are illustrated in Table 1.

Table 1. The perfect and the progressive (from Pratas 2018a: 109)

	Temporal ordering	Temporal ordering	Temporal meaning
bare verb (perfect-like reading)	T-Sit precedes TT	TU and TT coincide	Present Perfect
		TU is after TT	Past Perfect
progressive morphology	T-Sit includes/ coincides with TT	TU and TT coincide	Present progressive
		TU is after TT	Past progressive

These aspectual distinctions will be inserted into a more complicated picture below, which includes the more salient variation regarding other functional items (Pratas 2018a,b).

3.2 The relevant temporal morphemes

This section describes the relevant temporal morphemes in active clauses and the corresponding meanings in both varieties.

Summarizing the data which are common to both varieties, with only some minor morphophonological differences, we have the present progressive and the present perfect, as in (1), above, repeated here as (3) for convenience, but now with a more direct translation for the perfect-like reading.

- (3) a. *N sata/tita bebi/bebê agu.*
 1SG SATA/TITA drink water
 ‘‘I am drinking water.’’
- b. *N bebi/bebê agu.*
 1SG drink water
 ‘‘I have drunk (the) water.’’

Also common to both varieties are the present attitudinal / present habitual / future, which are marked with preverbal *ta* (may be *te* in São Vicente), as illustrated in (4).⁸

8. For a detailed description of more temporal morphemes in São Vicente, see Swolkien (2015).

- (4) *N ta dá aula di portuges.*
 1SG TA give class of Portuguese
 “I am a Portuguese teacher.”
 “I teach Portuguese.”
 “I will give a Portuguese class.” (e.g., if preceded by an adverb, like *tomorrow*)

When something in the linguistic context orders TU after TT, all these values exhibit morphological differences between the two varieties, including some distinct position of the morphemes regarding the verb. This is the case of the perfect in the past, which in Santiago is *V-ba* and in São Vicente is *tinha* (suppletive past form of Portuguese *ter* “have”) + participle, illustrated in (5) and (6).

- (5) [...] *N bebeba agu.*
 [...] 1SG drink:BA water
 “[before I went to bed] I had drunk (the) water.”
- (6) [...] *N tinha kmid pex.*
 [...] 1SG had eaten fish
 “[before I went to bed] I had eaten (the) fish.”

This is also the case with the attitudinal/habitual in the past. Depending on the context, the same configurations may have a conditional interpretation. In Santiago, we have these values marked with *ta V-ba*, and in São Vicente with *tava V* (with its own variant *tá V*), illustrated in (7) and (8).

- (7) *N ta viveba ma nha madrasta i nha pai [...]*
 1SG TA live:BA with my stepmother and my father [...]
 “I used to/would live with my stepmother and my father [...].”
- (8) *N tava vivê ma nha madrasta i nha pai [...]*
 1SG TAVA live with my stepmother and my father [...]
 “I used to/would live with my stepmother and my father [...].”

Finally, this is the case also with the progressive in the past, which is marked with *sata V-ba* in Santiago, and with *tava* or *tá + ta* or *te V* in São Vicente. We have seen this in the simple sentences in (2) and confirm it with the complex sentences in (9) and (10).

- (9) *Kantu bu txiga N sata trabadjaba.*
 When 2SG arrive 1SG SATA work:BA
 “When you arrived I was working.”
- (10) *Kond João txegá, Ana el inda tava ta trabaia.*
 When João arrive, Ana 3SG still TAVA TA work
 “When João arrived, Ana was still working.”

Table 2 summarizes the correspondences between the different morpheme combinations and the temporal meanings of the clauses in which they appear.

Table 2. Temporal morphemes in both varieties (from Pratas 2018b)

	Santiago	São Vicente
zero/null	gives the bare form of some predicates what seems a past reading (1b)	gives the bare form of some predicates what seems a past reading (1b)
preverbal progressive	<i>sata/aita</i> used <u>in non-past and past</u> progressives ((1a) and (2a), (9))	<i>tital/tite</i> used <u>in non-past</u> progressives (1a)
preverbal <i>ta</i>	used <u>in non-past and past</u> habituals, attitudinal, generics, futures ((4), (7))	may also be <i>te</i> ; used <u>in non-past</u> habituals, attitudinal, generics, subsequents (4)
postverbal <i>-ba</i>	combines with the above morphemes and appears associated <u>with past</u> situations ((5), (9), (7), respectively)	-----
preverbal <i>tava/tá</i>	-----	used <u>in past</u> habituals, attitudinal, generics and subsequents (8) combines with <i>ta</i> or <i>te</i> <u>in past</u> progressives (10)

Given that past tense is defined, as above, as the ordering of the Time of Utterance after the Topic Time (Klein 1994, 2010), Santiago *-ba* looks pretty much like a tense marker – a position defended in Pratas (2007, 2012, 2014). The contribution of São Vicente preverbal allomorphs *tava/tá* seems less straightforward, since it marks more than just the order of TU regarding TT. But at least the case of *-ba* had to be addressed, and the point in my recent works, also assumed here, is that there are reasons to defend a different view. The main argument for this change of perspective is this: Santiago *-ba* occurs in clauses where it has no regular past meaning, as shown in the following example in bold. Note that this example corresponds to nonfinite environments in languages with a clear finite/nonfinite distinction, such as English and Portuguese.

(11) *Otu algen ta fikaba lá na igrexa pa ka moreba.*

Other people TA stay:BA LOC PREP church PREP NEG die:BA

lit. “[when there were floods] other people would take refuge in the church, to not die.” (so that they would not die)

Moreover, in conditionals *-ba* may associate with preverbal *ta* and similarly not convey an obligatory past tense. This is illustrated in (12), from Pratas (2007: 72), where neither the main clause nor the conditional one has a strict past interpretation:

- (12) *Si N fikaba un anu na Kauberdi, N ta papiaba kriolu dretu.*
 If 1SG stay.BA one year in Cabo Verde, 1SG TA speak.BA Kriolu right
 “If I stayed for a year in Cabo Verde, I would speak Creole perfectly.”

At this point a description of the language’s temporal readings under modals may help us clarify the temporal readings of these morphemes in more complex sentences, and this is the topic of the next subsection.

3.2.1 *Temporal readings under modals*

The following data were partly included in Pratas & van de Vate (2012) with a different goal: we wanted to compare the sentence structure of epistemic vs. deontic modals in Capeverdean and Saamáka. Here, they specifically illustrate the temporal meanings of these embedded clauses, to which some other examples have been added.

Following Condoravdi (2002) and Laca (manuscript), a modal clause has two time intervals: (i) a temporal perspective, or the modal anchor time; and (ii) a temporal orientation, which refers to “the time at which the temporal property is instantiated”, i.e., the modal evaluation time (Laca manuscript: 4).

In Table 3, we have the different relations between the modal base and the aspectual groups of the embedded predicate,⁹ and the temporal orientation in each case. An epistemic modal conveys the knowledge of the speaker about the situation denoted under it, while non-epistemic stands for other values, such as the circumstantial (the claim of the speaker given the circumstances of the world), the deontic (an expectation or desire of the speaker related to how the world should be), etc. – the different sub-types among those here called non-epistemic are not relevant to the current description.

Table 3. Temporal readings under modals (from Pratas 2018b)

	Non-epistemic	Epistemic
stative situations	subsequence	simultaneity
dynamic situations (activities, accomplishments and achievements)	subsequence	precedence

9. As noted in footnote 3, the event types assumed are the ones described in Vendler (1957). Activities, accomplishments and achievements are subsumed under the label ‘dynamic predicates’, since the differences among them are not relevant for the current purposes.

This is illustrated in the examples below, from Santiago.

Non-epistemic (permission/possibility and obligation, respectively)

- (13) *Bu pode sta diskansadu, N ta txiga sedu.*
 2SG MOD be rested, 1SG TA arrive early
 “You can rest easy, I will arrive early.”
 [non-epistemic + stative = subsequence]
- (14) *Bu ka debe fla kasi pa bo mosa.*
 2SG NEG MOD say/tell lie PREP your girl
 “You must not lie to your girlfriend.”
 [non-epistemic + dynamic = subsequence]

Note that the modal *pode* “can” in (13) is one of permission/possibility and, thus, it is bare. The same superficial verb, but with an ability reading, must be marked by *ta* in the present, which is a hint about its non-stative reading (like in *be able to pay the rent* – present reading: *ta pode* in both varieties; past reading: *ta podeba* in Santiago and *tava podê* in São Vicente).¹⁰

Epistemic

- (15) *E ka debe sta dretu di kabesa.*
 3SG NEG MOD BE well of head
 “S/he must be out of her/his mind.”
 [epistemic + stative = simultaneity]

In the above sentence we get a simultaneous interpretation for the modal and the stative situation denoted by the bare predicate below it.

In the next one there is a precedence reading of the relevant situation denoted by the bare dynamic predicate. Despite this, however, and given that this precedence is denoted by a bare verb, the moment of evaluation also coincides with the time of the modal – in other words, we again have a perfect-like interpretation.¹¹

- (16) Context: two friends meet in the evening and one of them wonders whether a third friend has worked on that holiday; knowing their friend’s habits, the second one says:
Luisa debe trabadja.
 Luisa MOD work
 “Luisa must have worked.”
 [epistemic + dynamic = precedence (of the relevant dynamic situation)]
 (adapted from Pratas & van de Vate 2012: 416)

10. I am thankful to Dominika Swolkien (p.c.) for calling my attention to this for São Vicente.

11. For a proposal that perhaps all situations in Capeverdean are either basic or derived states, see Pratas (2010: 229). Here, we would again have a state, since the temporal reading of perfects is related to the post-state of the situation itself, as described in § 3.1.

This also leads us to another fact about epistemic modals in the language: they occur in biclausal structures with the lower subject raising to the matrix subject position (the underlying meaning of these sentences is something like “It is possible/probable that ...”). We see that there is a biclausal structure because the embedded verb may be marked by the progressive morpheme.

- (17) *Ka bu faze raboliso pamodi Maria pode sata durmi.*
 NEG 2SG make noise because Maria MOD SATA sleep
 “Don’t you make any noise because Maria may be sleeping.”
 (adapted from Pratas & van de Vate 2012: 426)

Under non-epistemic modals the perfect-like is not available. This suggests that what we have in this case is indeed a monoclausal configuration, and the temporal orientation of subsequence depends on other semantic features of the modal – it establishes conditions that are to be met after TT (the modal anchor time).

Now it becomes even clearer – and this is the relevant point for current purposes – that any *-ba* occurring under the modal is not tense. In fact, the same relations hold when the temporal perspective changes, that is, when TU is after TT.

Non-epistemic (desire and obligation, respectively)

- (18) *N kreba serba veterinario, mas N bai pa profesor.*
 1SG want:BA be:BA vet, but 1SG go for professor
 “I wanted to be a veterinarian, but I ended up a teacher.”
 [non-epistemic + stative = subsequence] (Pratas 2007: 104)
- (19) *Nu ten ki lababa mon... pa nu trabadjaba na midikamentu.*
 1PL have that wash:BA hand... PREP 1PL work:BA PREP medicine
 “We had to wash our hands... to work with medicine.”
 [non-epistemic + dynamic = subsequence] (Pratas 2018b)

Epistemic

- (20) *N diskunfia ma kel omi debaba serba nha tio.*
 1SG guess that REL man mod:BA be:BA my uncle
 “I have guessed that that man should be my uncle.”
 [epistemic + stative = simultaneity] (Pratas 2018b)

Therefore, the rationale for the alternative proposal in Pratas (2018a; b) is the following. If the verbs embedded by non-epistemic modals do not allow the perfect-like reading – which proves that there is no direct relation between this embedded TT and the TU – there are absolutely no grounds to argue that this embedded *-ba* is a tense morpheme. It is then not difficult to sustain that an alternative analysis for this *-ba* may be extended to the one embedded under all kinds of modal constructions, and also to the one occurring in purpose adverbial clauses.

Since some of these embedded predicates with *-ba* have a modal meaning – nothing is said about whether, or when, these situations truly occurred – or may even have a counterfactual value, it could be argued that what we have here is a mood agreement morpheme (of the type proposed in Matthewson (2010: 2), about a clause in which a main verb of doubt licenses subjunctive mood in the lower clause, in an Italian example from Palmer (2006: 117)). It sometimes occurs in modal contexts whose embedding situations are not in the past, but even then the strict mood agreement hypothesis seems not correct either.

The proposal then was that *-ba* is a temporal agreement/concord¹² morpheme associated with some past environments. It appears (i) in root clauses where a past interpretation is provided by the context, and (ii) on subordinate verbs whose embedding clauses often convey past situations.

The same reasoning holds for the preverbal *tava / tá* in the São Vicente variety, with some specifics related to the diachrony of *ta*. In order to account for these different morphological forms – defending that they are a case of micro-variation in the language, in line with the Borer-Chomsky conjecture (Baker 2008) – another conjecture has been formulated: Capeverdean *ta* underwent a complete progressive cycle (Pratas 2018a: 120).

3.2.2 *A conjecture on the diachrony of these allomorphs*

The origin of *ta* is quite undisputed. It comes from the Portuguese reduced form of *está*, the third person singular of *estar* “be” (stage-level). As was documented in Schuchardt (1882),¹³ it was initially the progressive marker in Capeverdean, which then generalized its use to a more general imperfective meaning, including habituals (as was also proposed in Kihm 1994 for the Guinea-Bissau Creole).

My take on *ta* goes further, in that *ta* underwent a complete progressive cycle, such as the one proposed for progressive markers in other languages (see Dahl 1985; Bybee & Dahl 1989; Bybee, Perkins & Pagliuca 1994; Deo 2015, among others). I argue that after that shift of *ta* to a more general imperfective use referred by Schuchardt (1882: 911, translated in Hagemeyer & Holm 2008) for Capeverdean, and Kihm (1994) for Guinea-Bissau Creole, some reinforcement emerged to fill the gap of a more specific progressive marker (Pratas 2018a; b). The account for the progressive reading of *ta* in embedded clauses headed by perceptive verbs (*N odja Maria ta badja* “I have seen Maria dancing”) then nicely follows: here *ta* keeps its original

12. This double classification here is just to avoid any confusion with syntactic agreement.

13. In Schuchardt (1882: 911, translated in Hagemeyer & Holm 2008) there is the observation that, “originally”, Capeverdean *N ta dá* means, in Portuguese, *eu estou dando* or *eu estou a dar* “I am giving”, a meaning that has later “blurred to [eu dou]” “I give” (adapted from Hagemeyer & Holm 2008: 148).

progressive meaning, for there is no possible ambiguity with the habitual. As for the progressive in root clauses, we now have both the periphrastic forms where *ta* has been reinforced by locative auxiliaries, and the alternative non-periphrastic forms which always involve *ta* (or the allomorph *te*) preceded by some other morphemes with a locative content. Specifically, in Santiago we now have the progressive forms *sta+ta+V* and the ones with variants of the preverbal morpheme *sata* (*ata*, *aita*).

In past environments, licensed by some context that locates TU after TT, there is the temporal agreement/concord *-ba* on the verb; this *-ba* sometimes appears doubled in what seem nonfinite embedded predicates. In São Vicente, we now have the progressive form *tita* (or *tite*), and a periphrastic form with *stod* (*stod+ta+V*). In the transition from the southern islands, *-ba* was lost (Swolkien 2015: 233 fn. 82). By analogy with the Portuguese first and third persons singular, past imperfective of *estar* (*estava*), this Capeverdean *tava* is a multifunctional form related to distinct past environments:¹⁴

- i. in past habituals (*tava+V*);
- ii. as the reinforcement of past progressives (*tava+ta* or *te+V*).

As observed above, this is a case of micro-variation in which different morphological forms express the meaning of the same lexical item (although this meaning may be combined in different ways with other lexical items, whether they are other temporal morphemes or just the verb itself).

3.3 A temporal agreement marker in both varieties

Capeverdean displays the dual modal values described in Comrie (1985: 45): “... some languages have a basic modal distinction between realis and irrealis, where realis refers to situations that have actually taken place or are actually taking place, while irrealis is used for more hypothetical situations, including situations that represent inductive generalisations, and also predictions, including also predictions about the future.” This corresponds to what has been described above:

14. This form is also used as a copula, like the reduced form of the Portuguese *estava*:

- (i) *N tava un bokod grand ...*
1SG was one bit big
“I was a bit grown-up ...”

And then, when we need the same copula value modified by some specific aspect / mood, the copula itself is *stod*, as said above, and we get *tava* as the temporal marker just described:

- (ii) *Un ves, pais tava stod na kaza, tud na pé d fidj ...*
one time, parents TAVA be in house, all at foot of child
“In the old days, parents used to be home, together with their children ...”

- i. an irrealis mood, in its various values (generics, futures, attitudinals);
- ii. a realis mood, expressed through the perfect or the progressive.

Therefore, we are free to accept that in natural language past, present or future meanings of a sentence may be expressed without any dedicated tense morphemes. This occurs in this language: there is a distinction that is strictly of mood (realis vs. irrealis) and, within the realis mood, there is an aspectual distinction between the perfect and the progressive. To be shifted into a ‘past’ interpretation, all these irrealis and realis meanings need a context (discourse, adverbs, or other sources of temporal information) which locates TU after TT. In the absence of this context, an interpretation where TU and TT coincide is obtained. Therefore, the specific morphological forms associated with the past-shifted versions of these mood and aspect meanings are not tense markers, rather behaving like temporal agreement/concord: they do not bring any past meaning in themselves – in conditionals, for instance, they may even occur in the embedded clauses which refer to the present or the future (in cases similar to those which Iatridou (2000) has described as ‘fake past’ for Modern Greek).

As noted above, there is an underspecification in the mapping between the functional head at stake here and the morphological forms (Adger & Smith 2010), allowing the language to have the different units for this lexical item (temporal concord/agreement) that we find in each variety: postverbal *-ba* in the older variety of Santiago; in the younger variety of São Vicente, there is *tava/tá*, which is more complex since it incorporates either the habitual meaning or the locative reinforcement of the progressive. Also in São Vicente, some suppletive forms from Portuguese are used in other contexts where Santiago has *-ba*, such as *tinha* ‘had’+participle for the past perfect. The following tables (from Pratas 2018a: 119) illustrate these relations.

Table 4. Mood and aspect values, also with *-ba* (Santiago)

	Realis		Irrealis
	Perfect	Progressive	Habitual or prospective
TU coincides with TT	V (present perfect) (1b)	<i>sata</i> V (present progressive) (1a)	<i>ta</i> V (present habitual or future) (4)
TU is after TT (provided by the context)	V- <i>ba</i> (past perfect) (5)	<i>sata</i> V- <i>ba</i> (past progressive) (9)	<i>ta</i> V- <i>ba</i> (past habitual or conditional) (7)

Table 5. Mood and aspect values, also with *tava/tá* (São Vicente)

	Realis		Irrealis
	Perfect	Progressive	Habitual or prospective
TU coincides with TT	V (present perfect) (1b)	<i>tita/tite V</i> (present progressive) (1a)	<i>ta V</i> (present habitual or future) (4)
TU is after TT (provided by the context)	<i>tinha</i> + participle (past perfect) (6)	<i>tava</i> or <i>tá</i> + <i>ta</i> or <i>te V</i> (past progressive) (10)	<i>tava</i> or <i>tá V</i> (past habitual or conditional) (8)

This recent proposal, however, did not explicitly state what these morphemes agree with. This is the topic of the next section.

4. Marking other values

The notion of low accessibility that I started to explore in Pratas (2021) applies to both past habituais and past progressives (inaccessibility), on the one hand, and to conditionals and other typical modal contexts (several levels of low accessibility), on the other hand.

As pointed out in the introduction, this novel approach uses a different notion of accessibility from that used in Arregui (2005). As a parallel note, this notion is also different from the proposal in Cable (2011, 2013) on what he calls Temporal Remoteness Morphemes (TRMs) in the Bantu language Gikūyū (Kikuyu). The author argues that these are prefixes which occur in past-tensed verbs and provide further information about the distance between the event described and the time of speech. They possess the presuppositional semantics commonly hypothesized for tenses, but “some also possess an indexical semantics more characteristic of temporal adverbials. ... I propose that (like tenses), the TRMs of Kikuyu are temporal pronouns, but (unlike tenses) restrict the Event Time rather than the Topic Time” (Cable 2011: 3).

The notion of remoteness, in the sense of distancing, would also be of interest to account for the speaker’s attitude towards the situations denoted in the Capeverdean sentences under analysis here. But on the other hand it might be misleading: remoteness, when connected with what seems past morphology, might point exclusively to different degrees of temporal distance, and this is not the idea at all.

The idea here is this:

- The speaker identifies past situations (i.e., situations whose Topic Time is before the Time of Utterance, which is given by the context) with a sense of inaccessibility; the past is gone, and although we may access it in our memories we can never go there again; hence we may say it has a quite low accessibility.
- There is a morpheme associated with these past constructions, just because it expresses this concept of low accessibility – note that what conveys this notion is the speaker’s knowledge of the world and their perceptions about time, not any morphology of the past.
- There are other situations that, although not necessarily in the past, also suffer from a low degree of accessibility under the speaker’s perspective, and therefore they use the same linguistic forms as above to convey this low accessibility concept.
- This is how we get conditionals (which are not necessarily in the past, but point to a time viewed as lowly accessible all the same); and this is how we get desires and wishes also marked by this same morphology – because the speaker intends to convey this low accessibility meaning, such as in *N tava gostá / N ta gostaba di bai sinema* “I would like to go to the movies”, meaning “I want to” but with a reduced amount of hope regarding the success of this wish.

This proposal, which implies disentangling a past tense meaning from the temporal markers sometimes – not always – associated with it, gains even more strength with the observation that this relation between a past meaning and a low level of accessibility produces other interesting results: when the speaker wants to make past situations look more accessible, they narrate them in the present (the so-called narrative ‘present’), which helps them disguise that mismatch between the Time of Utterance and the Topic Time. They know that the narrated situations occurred in the past (this is the meaning level) but they tell them as if they were not marked by low accessibility (this is the morphology level). This type of narrative ‘present’ (among the “non-canonical usages” of the present described in Klein 2010: 48–51) is attested in many languages, and perhaps the strategy is of the same sort in all of them (this is a case for future study).

The relation between the levels involved in this computation is also interesting in that the present perfect morphology, which, as said in § 3, denotes situations which occurred prior to the Time of Utterance (the TT and TU coincide, but the T-Sit precedes TT), is never involved in these low accessibility contexts. This shows that what truly makes the difference regarding these uses in the the past (which is no longer fully accessible, and this is expressed morphologically, just like other inaccessible situations are expressed in a similar way even if they are not past) is

the precedence of the Topic Time regarding the Time of Utterance. In other words, what counts to build this notion is the TU's relation not with the T-Sit but rather with the TT.

Another novel characteristic of this proposal is that it puts together constructions which are easily taken as having a modal meaning and others which have been previously treated under a tense analysis.

One way of putting this analysis to the test is to check whether it works also for other temporal morphemes in other languages. And this is indeed the case with Portuguese past imperfective markers, which are used in past and in non-past temporal meanings, in the same way these Capeverdean morphemes do. Crucially, both Capeverdean allomorphs under study here are derived from one of the regular past imperfective morphemes in Portuguese, *-va* (as illustrated in (21b); the other common form is *-ia*, as illustrated in (21a)). Nevertheless, contexts such as the ones below have never, to my knowledge, been analysed under this perspective – that these morphemes are not any form of past markers which sometimes are used as 'fake pasts', but rather markers for an entirely distinct layer of temporal meaning.

- (21) a. *Eu comia agora um gelado.*
 I eat:LAM now an ice-cream
 "I would now eat an ice-cream."
- b. *Amanhã, se estivesse calor, eu almoçava na esplanada.*
 tomorrow, if be:SUBJ heat, I lunch:LAM at terrace
 "If it were hot tomorrow, I would lunch alfresco."

The future study of these low accessibility morphemes (LAMs) will also consider other works on different temporal constructions that are used by the speakers to provide similar temporal meanings, although they may not have been identified as such.

5. Conclusion and a note regarding future research

In this paper I further confirm a new perspective on the expression of temporal meaning in Capeverdean which accounts for the use of apparently past morphology in contexts where a past meaning is not guaranteed. My proposal is that in both cases – the ones with a past interpretation and the ones that do not have it – this same morphology does not mark past, but rather a different level of meaning: a low degree of accessibility from the speaker's perspective. This takes as a point of departure the fact that in Capeverdean a past meaning occurs without any dedicated tense morpheme: there is a distinction that is strictly of mood (realis vs. irrealis) and,

within the realis mood, there is an aspectual distinction between the perfect and the progressive. To be shifted into a past interpretation, all these irrealis and realis meanings need a context (discourse, adverbs, other linguistic information, etc.) which locates TU after TT; in the absence of this context, a present interpretation is obtained through a pragmatic inference. The specific morphological forms associated with the past-shifted versions of these mood and aspect meanings thus behave like temporal agreement/concord rather than strict tense markers. The novelty here is that they agree, not with tense, but with that low accessibility notion which is a particular feature of some time locations – the past, certainly, but other unspecific locations as well. Therefore, both semantic content (a past meaning, provided by the context) and pragmatic inferences (principles about the functioning of the real world) are at stake in this computation. The syntactic implementation of this is left for future research, which must involve a projection for these LAMs.

Furthermore, there is an underspecification in the mapping between this functional item (temporal agreement/concord) and the morphological forms (Adger & Smith 2010), allowing the language to have the different units we find in each variety: in the older variety of Santiago, we have the postverbal *-ba*; in the younger variety of São Vicente, we have *tava/tá*, which is more complex, since it incorporates this temporal agreement and either the habitual meaning or the locative reinforcement of the progressive. Also in São Vicente, some suppletive forms from Portuguese are used in other contexts where Santiago has *-ba*, such as *tinha* + the participle for the past perfect.

Acknowledgements

I want to thank all my consultants in Cabo Verde and in Lisbon, the audiences in the workshops *Tenselessness* (Greenwich, London, October 2017) and *GREAT*, within *Going Romance 32* (Utrecht, December 2018), and two anonymous reviewers for their helpful comments and suggestions.

Funding

This work has been sponsored by national funding through FCT, *Fundação para a Ciência e a Tecnologia*, IP, within the scope of *LUDVIC – Language Unity and Diversity: Variation in Capeverdean and beyond* (IF/00066/2015), an exploratory research project associated with my current position as Principal Investigator.

References

- Aboh, Enoch & Michel DeGraff. 2017. A null theory of creole formation based on Universal Grammar. In Ian Roberts (ed.), *The Oxford handbook of Universal Grammar*, 401–458. Oxford: Oxford University Press. <https://doi.org/10.1093/oxfordhb/9780199573776.013.18>
- Adger, David & Jennifer Smith. 2010. Variation in agreement: A lexical feature-based approach. *Lingua* 120. 1109–1134. <https://doi.org/10.1016/j.lingua.2008.05.007>
- Arregui, Ana. 2005. On the accessibility of possible worlds: The role of tense and aspect. Ph.D. Dissertation: University of Massachusetts Amherst.
- Baker, Mark. 2008. The macroparameter in a microparametric world. In Theresa Biberauer (ed.), *The limits of variation*, 351–373. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.132.16bak>
- Baptista, Marlyse. 2002. *The syntax of Cape Verdean Creole, the Sotavento varieties*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.54>
- Bertinetto, Pier Marco, Karen H. Ebert & Casper de Groot. 2000. The progressive in Europe. In Östen Dahl (ed.), *Tense and aspect in the languages of Europe*, 517–558. Berlin: De Gruyter.
- Binnick, Robert I. 2010. Review of Kasia M. Jaszczolt, *Representing time: An essay on temporality as modality*. Oxford: Oxford University Press, 2009. *Journal of Linguistics* 46. 514–518. <https://doi.org/10.1017/S0022226710000125>
- Bybee, Joan L. & Östen Dahl. 1989. The creation of tense and aspect systems in the languages of the world, *Studies in Language* 13. 51–103. <https://doi.org/10.1075/sl.13.1.03byb>
- Bybee, Joan L., Revere D. Perkins & William Pagliuca. 1994. *The evolution of grammar: Tense, aspect and modality in the languages of the world*. Chicago: University of Chicago Press.
- Cable, Seth. 2011. Remoteness distinctions in the tense-aspect system of Kikuyu. lingbuzz/001374/current.pdf
- Cable, Seth. 2013. Beyond the past, present, and future: Towards the semantics of ‘graded tense’ in Gikūyū. *Natural Language Semantics* 21. 219–276. <https://doi.org/10.1007/s11050-012-9092-3>
- Comrie, Bernard. 1985. *Tense*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139165815>
- Condoravdi, Cleo. 2002. Temporal interpretation of modals: Modals for the present and for the past. In David I. Beaver, Luis D. Casillas Martínez, Brady Z. Clark & Stefan Kaufmann (eds.), *The construction of meaning*, 59–88. Stanford, CA: CSLI.
- Dahl, Östen. 1985. *Tense and aspect systems*. Oxford: Blackwell.
- Dahl, Östen & Viveka Velupillai. 2005. Tense and aspect. In Martin Haspelmath, Matthew S. Dryer, David Gil & Bernard Comrie (eds.), *The world atlas of language structures*, 266–281. Oxford: Oxford University Press.
- Demirdache, Hamida, & Myriam Uribe-Etxebarria. 2000. The primitives of temporal relations. In Roger Martin (ed.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 157–186. Cambridge, MA: MIT Press.
- Deo, Ashwini. 2015. The semantic and pragmatic underpinnings of grammaticalization paths: The progressive to imperfective shift. *Semantics & Pragmatics* 8. 1–52. <https://doi.org/10.3765/sp.8.14>
- Gillier, Raïssa. 2019. Orografia e identidade em cabo-verdiano. Paper presented at the 19th Annual Conference of the Association of Portuguese and Spanish-lexified Creoles & Summer Conference of the Society for Pidgin and Creole Linguistics, Lisbon, June 17–19.

- Hagemeyer, Tjerk & John Holm. 2008. Schuchardt, Hugo. On the Creole Portuguese of São Tomé (West Africa). Annotated translation from the German of 'Ueber das Negerportugiesische von S. Thomé (Westafrika)' *Sitzungsberichte der kaiserlichen Akademie der Wissenschaften zu Wien* 101(2): 889–917 [1882]. In John Holm & Susanne Michaelis (eds.), *Contact languages: Critical concepts in linguistics*, 131–156. London: Routledge.
- Iatridou, Sabine. 2000. The grammatical ingredients of counterfactuality. *Linguistic Inquiry* 31. 231–270. <https://doi.org/10.1162/002438900554352>
- Kihm, Alain. 1994. *Kriyol syntax: The Portuguese-based creole language of Guinea-Bissau*. Amsterdam: John Benjamins. <https://doi.org/10.1075/cll.14>
- Klein, Wolfgang. 1994. *Time in language*. London: Routledge.
- Klein, Wolfgang. 2010. On times and arguments. *Linguistics* 48. 1221–1253. <https://doi.org/10.1515/ling.2010.040>
- Klein, Wolfgang. 2014. Is aspect time-relational? Commentary on the paper by Jürgen Bohemeyer. *Natural Language & Linguistic Theory* 32(3). 955–971. <https://doi.org/10.1007/s11049-014-9240-1>
- Labov, William. 1969. Contraction, deletion, and inherent variability of the English copula. *Language* 45(4). 715–762. <https://doi.org/10.2307/412333>
- Laca, Brenda. Manuscript. *On modal tenses and tensed modals*. Université Paris 8. 2008.
- Lang, Jürgen. 2014. *A variação geográfica do Crioulo Caboverdiano*. Erlangen: FAU University Press.
- Matthewson, Lisa. 2004. On the methodology of semantic fieldwork. *International Journal of American Linguistics* 70. 369–415. <https://doi.org/10.1086/429207>
- Matthewson, Lisa. 2010. Cross-linguistic variation in modality systems: The role of mood. *Semantics & Pragmatics* 3. 1–74. <https://doi.org/10.3765/sp.3.9>
- Moens, Marc & Mark Steedman. 1988. Temporal ontology and temporal reference. *Computational Linguistics* 14(2). 15–28.
- Mufwene, Salikoko S. 2010. *Language evolution: Contact, competition, and change*. London: Continuum Press. <https://doi.org/10.1017/S0047404510000096>
- Palmer, Frank Robert. 2006. *Mood and modality*. Cambridge: Cambridge University Press.
- Portner, Paul. 2011. Perfect and progressive. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, 1217–1261. Berlin: Mouton de Gruyter.
- Pratas, Fernanda. 2007. Tense features and argument structure in Capeverdean predicates. Ph.D. Dissertation: Universidade Nova de Lisboa.
- Pratas, Fernanda. 2010. States and temporal interpretation in Capeverdean. In Reineke Bok-Bennema, Brigitte Kampers-Manhe & Bart Hollebrandse (eds.), *Romance languages and linguistic theory 2008*, 215–231. Amsterdam: John Benjamins. <https://doi.org/10.1075/rllt.2.12pra>
- Pratas, Fernanda. 2012. 'I know the answer': A Perfect State in Capeverdean. In Irene Franco, Sara Lusini & Andrés Saab (eds.), *Romance languages and linguistic theory 2010*, 65–86. Amsterdam: John Benjamins. <https://doi.org/10.1075/rllt.4.04pra>
- Pratas, Fernanda. 2014. The Perfective, the Progressive and the (dis)closure of situations: Comment on the paper by María J. Arche. *Natural Language & Linguistic Theory* 32(3). 833–853. <https://doi.org/10.1007/s11049-014-9236-x>
- Pratas, Fernanda. 2018a. Progressive forms and meanings: The curious case of Capeverdean. *Estudos de Lingüística Galega* 10. 103–128. <https://doi.org/10.15304/elg.10.5045>

- Pratas, Fernanda. 2018b. Tense results from mood and aspect: Temporal meanings in Capeverdean. Paper presented at GREAT, within Going Romance 32, Utrecht, December 10–14.
- Pratas, Fernanda. 2021. *The expression of temporal meaning in Caboverdean*. Berlin: De Gruyter. <https://doi.org/10.1515/9783110626629>
- Pratas, Fernanda & Marleen van de Vate. 2012. *Proceedings of the Annual Meeting of the Berkeley Linguistics Society* 38, 415–430. Linguistic Society of America. <https://doi.org/10.3765/bls.v38i0.3344>
- Reichenbach, Hans. 1947. The tenses of verbs. Section 51 of *Elements of Symbolic Logic*, 287–98. New York: The Macmillan Company.
- Silva, Izione Santos. 1985. Variation and change in the verbal system of Capeverdean Crioulo. Ph.D. Dissertation: University of Michigan.
- Silva, Izione Santos. 1990. Tense and aspect in Capeverdean Crioulo. In John Victor Singler (ed.), *Pidgin and creole tense-mood-aspect systems*, 143–168. Amsterdam: John Benjamins. <https://doi.org/10.1075/cll.6.06sil>
- Smith, Carlota. 1991. *The parameter of aspect*. New York: Kluwer. <https://doi.org/10.1007/978-94-015-7911-7>
- Stowell, Tim. 2014. Capturing simultaneity: A commentary on the paper by Hamida Demirdache and Myriam Uribe-Etxebarria. *Natural Language & Linguistic Theory* 32(3). 897–915. <https://doi.org/10.1007/s11049-014-9241-0>
- Suzuki, Miki. Manuscript. *The markers in Cape Verdean CP*. CUNY. 1994.
- Swolkien, Dominika. 2015. The Cape Verdean Creole of São Vicente: Its genesis and structure. Ph.D. Dissertation: Universidade de Coimbra.
- Swolkien, Dominika & Alexander Cobbinah. 2019. Cape Verdean Creole – Santo Antão: What we know so far. *Journal of Ibero-Romance Creoles* 9(1). 145–175.
- Veiga, Manuel. 1995. *O Crioulo de Cabo Verde: Introdução à Gramática*. Praia: Instituto Nacional do Livro e do Disco, Instituto Nacional da Cultura.
- Vendler, Zeno. 1957. Verbs and times. *The Philosophical Review* 66(2). 143–160. <https://doi.org/10.2307/2182371>

Very extracted

On old Italian *molto*

Silvia Rossi and Cecilia Poletto

Goethe-Universität Frankfurt am Main

In this paper we consider cases of extraction of the degree word *molto* “very, much” from its modifier position within an AdjP (Giusti 2010a; b; Poletto 2014) in Old Italian. Such cases are reminiscent of Left Branch Extractions (LBE; Ross 1967), but, differently from what happens with adverb extraction in Slavic (Talić 2017), the conditions under which *molto*-extraction is possible are very restricted: *molto* can be extracted only when the AdjP is (or modifies a nominal expression) in post-copular predicative position but not outside a fully-fledged DP. We propose that the reason why the structure is so restricted has to do with the presence/absence of a phase boundary, while the reason why this has been lost in modern Italian is the loss of the target position in the CP layer. This allows us to show that the loss of the Verb Second property (V2) in Italian has fine-grained consequences in unexpected domains like quantifier extraction.

Keywords: degree words, old Italian, modifier extraction, left periphery

1. Introduction

Research on the evolution of the Old Romance languages out of Latin has concentrated up to now on phenomena like verb second, null subjects, placement of clitics, but has not considered other changes, which might prima facie look like simple lexical changes and therefore uninteresting for syntactic research. One such case is the fact that in Old Italian (OI), in the written Old Florentine variety from the 13th and the 14th century (see Salvi & Renzi 2010: 7), the system of quantifiers has undergone subtle modifications, like the substitution of forms (see the substitution of n-words like *neuno* with *nessuno*, the rise of quantifiers of the *qualche* type, the loss of *cosa* “thing” as an NPI). Within this set of apparently only lexical shifts, one quantifier has been surprisingly resilient in standard Italian,

namely *molto* “much/very”. This is particularly notable, since this Q word derived from the Latin quantity adjective *multus* has been substituted in Romance by various types of elements with distinct etymological origins ranging from nouns originally designating big objects (mountain, sack, world, boat, etc.) to gradable adjectives (thick, good) to adverbial forms in many other Romance varieties (cf. for instance French, Carlier 2011, and many Italo-Romance varieties, Rohlfs 1969: § 954; Poletto 2018). However, once we look at its behavior, *molto* displays interesting syntactic properties, which can help us better define the complex change related to the internal structure of nominal expressions and to the rise of articles as the lexical spell out of the most external layer of the DP phase.

The degree Q *molto* was and is still used as a simple adjective with the meaning “numerous” (1a), as an indefinite determiner (1b), and as a verbal (1c) and an adjectival modifier (1d) – a behavior which is typical of quantifiers developing out of Latin quantity adjectives (Camus Bergareche & Pérez-Saldanya 2011 on Old Catalan; see also De Clercq 2017 on the internal syntax of degree words).

- (1) a. *le molte scelte nel menù* (ModI)
 the.PL many.FPL choices.F in=the menu
 b. *Ho parlato con molti studenti*
 have.1SG talked with many.MPL students.PL
 “I talked to my students.”
 c. *Ho mangiato molto*
 have.1SG eaten much
 “I ate a lot”
 d. *una ragazza molto bella*
 a.FSG girl.FSG very beautiful.FSG
 “A very beautiful girl”

However, OI presented peculiar cases which really stand out to the modern eye because they are sharply ungrammatical in modern Italian: *molto* in Old Italian could escape out of its modifier position, move to a Left Peripheral sentence position and ‘strand’ the AdjP (alone or within a nominal expression), the AdvP, or (much more rarely and in later texts) the PP it modified.¹ We dub this phenomenon *molto*_{EX}.

1. Cases of *molto* extracted out of a PP are rather limited, just 23 out of the entire corpus (5 from Boccaccio). Most cases are like the one above where the PP is a sort of manner adverbial, “out of the straight way” meaning “being immoral”.

- (2) a. *La gente si meravigliò molto; e quelli cominciò questa sua*
 the people refl=wondered much and that.one started this his
canzonetta tanto soavemente quanto seppe il meglio, ché
 little.song so.much softly as.much knew the best, because
molto il sapea bene fare.
 very it=could well do
 “The people around were very impressed; and he started his chansonette
 as sweetly as he could possibly do, as he could do it very well”
 (Nov., p. 64)
- b. *Molto m'è dolce e soave il pensiero de li amici passati*
 Very me=is sweet and gentle the thought of the friends past
dal secolo;
 from.the century
 “The thought of the friends who passed away is very sweet and soothing
 to me”
 (FSI, p. 193)
- c. *Molto sareste fuori della diritta via dell'amore ...*
 very were.2PL out of=the straight way of.the love
 “You would be very much out of the right way of love and ...”
 (De Amore, L. 1 cap. 17)

The reason why *molto*_{EX} is theoretically interesting is that these cases are very much reminiscent of violations of the Left Branch Condition (LBE Ross 1967). In particular, Ross observed that English resists subextraction of degree words modifying adjectives: *How you are how tall! (see Corver 2017 for a recent overview). Ross already noted, however, that languages like Latin and Russian allowed LBE, and indeed, LBE of degree and adverb modifiers is allowed in Slavic (Talić 2017). Furthermore, Ross pointed out that languages allowing subextractions from Left Branches were either morphologically rich or had ‘scrambling’.

Yet, we cannot define OI as a language with a morphologically rich nominal declension, since it had already lost case declension in regular DPs, and the inflection in the nominal system is not different from ModIt, where these cases are all sharply ungrammatical. Since the direct relation to morphology does not seem a viable solution to explain the reason why *molto* can be extracted in OI but not in ModI, we surmise the reason is purely syntactic. We propose that the reason extraction out of a left branch is possible in OI has to do with a well-known change in the history of Italian, namely the change in the left periphery. As Poletto (2014) has shown, the left periphery of the DP and of the CP phases allows for movements in OI that are blocked in the modern language for independent reasons. We propose therefore that the loss of *molto*_{EX} is related to the loss of the V2 property for the CP and the same type of change in the determiner system for the DP, i.e., loss of N to

d movement and obligatory presence of a determiner. In the present contribution, we consider all cases of *molto* (mostly with the meaning of the English *very*) which are displaced or extracted as in the cases in (1). Our data are taken from all Old Florentine prose texts (including translations) from 1200 to 1370. The editions are those of the *Opera del Vocabolario Italiano* (OVI) database (<https://artfl-project.uchicago.edu/content/ovi>), and our set of examples has been put together through both specific queries of the database and extensive reading of the texts.²

2. The data: When and where is extraction possible?

We can split *molto*_{EX} into several subtypes with partially different properties.

- A. The first type is the one already illustrated in (2a) where *molto* is extracted out of a low IP adverb (we adopt here Cinque's (1999) definitions). Among the low adverbs, only those which are gradable can be modified by *molto* for semantic reasons. Therefore, the only adverbs found with *molto*_{EX} are: *ben(e)/mal(e)* "well/badly", *volentieri* "willingly", *meglio* "better", *meno* "less", and *più* "more" in comparative structures. Other types of adverbs, most notably gradable manner adverbs in *-mente* that are frequently modified by *molto*, are rarely found with *molto*_{EX}, though these cases are possible (we have found 9 instances, 7 from translations of original Old French texts). However, this should not be surprising: following Chomsky (2005) and in light of what we argue below, sub-extraction takes place through an intermediate step where the XP lower than the left branch first moves to a higher position leaving the left branch as the only lexical element in the remnant. Taking [*molto-Adv*] as the base order, the intermediate step for *molto*_{EX} to apply should be [*Adv-molto*], which however is virtually non-existent in OI, hence the surprising scarcity of *molto*_{EX} with *-mente* adverbs. Despite its limitation to a certain class of adverbs and the fact that this type is rather rare, when *molto*_{EX} modifies adverbs it is not sensitive to the verb and is found essentially with all verb types.
- B. The second type is extraction of *molto* out of an adjective in predicative position as the one in (2b), which is restricted to cases where *molto*_{EX} is extracted out of the predicative position of a copular sentence (as noted also in Giusti 2010b: 598 and Poletto 2014: 84ff.).

2. We have retrieved in total 323 cases of *molto*_E up to 1370. From these, we excluded 48 examples from Boccaccio, as it can be easily shown that this author exhibits a series of morphosyntactic idiosyncrasies that are not found in other texts. Space prevents us from fully investigating Boccaccio's syntax, which however certainly deserves to be thoroughly considered. The editions and abbreviations for the texts used here are listed in the Sources section at the end of the paper.

- (3) a. ... *che molto sarebbe lungo a contare e* ...
 ... that very would.be long to tell.inf and...
 “... that it would be very long to report ...” (VeV, cap. 50)
- b. *Molto è bella, ma la mia è assai più bella*»
 very is beautiful but the mine is much more beautiful
 “It is very beautiful but mine is much more beautiful.” (Nov., p. 84)

This also extends to cases like the following, where the Adj is still predicative but not in copular constructions:

- (4) a. *Et certo molto mi pare verisimile: ...*
 and certainly very to.me=looks likely
 “And it certainly looks to me very likely ...” (Rettorica, p. 29)
- b. *E quelli della casa de' Cavalcanti molto se ne*
 And those of=the house of.the C. very refl=of.it
mostrarono favorevoli.
 showed favourable
 “And the Cavalcanti people showed themselves very inclined to it.”
 (DCompagni, L. 3 cap. 6)

The empirical generalization that puts (3) and (4) together is that *molto* can be extracted out of an AdjP only when this is part of a small clause structure.

Molto can also be extracted out of a predicative nominal expression, which is evidently not a referential DP even in those cases in which we see a definite article:³

- (5) a. *ma molto è piccola cosa dire dell'arte sì come fece elli, e*
 but very is small thing talk.INF of.the art so how did he and
molto è grandissima parlare per l' arte, ...
 very is very.big talk.INF for the art
 “but it is a very insignificant thing to talk about art as he did, while it is a
 very excellent thing to talk for the art” (Rettorica, pag. 70)
- b. *Perché molto è grande iniquità la vita, che ...*
 because very is big iniquity the life, that
 “Because it is a very big iniquity the life that ...” (FR, cap. 39)

Note that OI admitted quite a lot of preposing to the DP initial position of various DP-internal elements (Poletto 2014: Chapter 3), which were eventually lost in the later stages of the language.

Concentrating on scrambling with *molto*, OI presented cases like the following (Giorgi 2010: 290; Giusti 2010b: 598; Poletto 2014: § 3.3.3) where *molto* modifies

3. We adopt here Longobardi's (1994) idea that D is the locus where referentiality is marked. Here no determiner is present.

a prenominal adjective, and cases in which the adjective has moved to a prenominal position stranding *molto* in the postnominal one which are ungrammatical in ModIt:

- (6) a. *Democrito fue molto grande filosofo*
 Democritus was very big philosopher
 “Democritus was a very important philosopher” (FSI, p. 106)
- b. *Allora trovò una molto bella canzonetta*
 Then found.3SG a very nice little.song
 “He then found a very nice chansonette” (Nov., p. 64 r 86)
- (7) *una donna giovane e di gentile aspetto molto...*
 a woman young and of genteel aspect very...
 “... a young woman with very noble looks ...” (VN, cap. 8, par. 1)

Thus, in light of these cases where the adjective can move independently from its modifier, it is tempting to analyse *molto*_{EX} in (2) to (5) as cases in which the Adjective has moved out of its phrase. Remnant movement of *molto* to a high Spec in the nominal structure follows, and from there *molto* can escape into a higher sentential projection.⁴ The same analysis has been proposed by various authors (see among others Abels 2003; Leu 2008) for cases of so-called ‘was für split’ in German. The derivation we propose is the following, where *molto* is in DegP sitting inside an AdjP, which is in turn in an FP within the nominal spine. For purposes of readability, we abstract away from N movement, which is not relevant to our point:

- (8) → merge *molto* in the DegP of the AdjP sitting in SpecF:
 a. [FP [SpecXP [DegP *molto* [AdjP A° *piccola*]]] [F°][NP [N° *cosa*]]]
 → move AdjP to the highest Spec of the XP containing it

4. An anonymous reviewer asks whether and how this can be related to cases of ‘scrambling’ of determiner *molto* as the following:

- (i) *Poi vidi cose dubitose molte* (VN 23.23, vv. 43)
 then saw.1SG things.F dubious.FPL many.F
 “Then I saw many frightening things” (Or even “Then I saw very frightening things.”)

Such cases are reported in Giusti (2010a: 394–395), who however observes that these cases are rather infrequent. In the specific case in (i), it is difficult to be certain whether *molte* is an actual postnominal D or a case of a scalar modifier agreeing with its modified adjective *dubitose* (such cases are very rare, see Giusti 2010b: 614). It should also be noticed that this example is in the verse part of Dante’s *Vita Nuova* and that *molte* rhymes with *discolte* a couple of lines below. More generally, it is often difficult to identify whether *molto* is a determiner or an adjective meaning “numerous”, and as an adjective it could appear post-nominally.

- b. [FP [XP [AdjP A° *piccola*] [DegP *molto* {_{AdjP} A° *piccola*}]] [F°] [NP [N° *cosa*]]]
 → move the remnant DegP containing *molto* and the trace of the Adj to the highest FP
- c. [FP [DegP *molto* {_{AdjP} A° *piccola*}] [FP [XP [AdjP A° *piccola*] {_{DegP} *molto* {_{AdjP} A° *piccola*}}]] [F°] [NP [N° *cosa*]]]
 → move the remnant containing *molto* and the trace of the Adj to the CP
 [CP ... [DegP *molto* {_{AdjP} A° *piccola*}].. [TP è ... [PRedP [FP [DegP *molto* {_{AdjP} A° *piccola*}] [d°] [FP [XP [AdjP A° *piccola*] {_{DegP} *molto* {_{AdjP} A° *piccola*}}]] [F°] [NP [N° *cosa*]]]]]]]

Predicative nominal expressions do not have a D layer, an assumption already made by Longobardi (1994) that is crucial to restrict extraction only to predicative nominal expressions. However, this does not mean that they do not have a functional spine of FPs, which is used for extraction.

As mentioned in Poletto (2014: 147), in OI cases of split-quantification with *molto* like (9a), (cf. the French *beaucoup*, Obenauer 1994; Baunaz 2011: 2), and cases where *molto* is extracted out of normal argumental/adjunct DP, cf. (9b) are unattested in the whole corpus.⁵

5. The OI pattern is all the more interesting compared to Old French. Carlier (2011) shows that, although French *beaucoup* appeared in the Middle French period, from the 15th cent. onwards, Old French had cases of ‘split’-quantification with *mult* since its earliest stages (Carlier 2011: 11 Examples (45)):

- (i) a. *Mult unt oïd e peines e ahans.* (Roland v. 267)
 Much have had and sorrows and torments
 “They had many sorrow and torments”
- b. *Mult sunt de malvais estres* (Philippe de Thaon, *Bestiary*)
 Much are of bad creatures
 “There are many bad creatures.”

Old French did have cases of *mult*_{EX} parallel to those we find in OI (ii), but it could also have *mult* extracted out of an adjective in what looks like a regular DP (that is, we do find cases like (9) above, cf. (iii) (Carlier 2011: 14 Example (58)).

- (ii) *Tere de France mult estes dulz país* (Roland v. 1861)
 Land of France much/very are.you sweet country
 “Land of France, you are a very sweet country”
- (iii) *Entres les dous oilz mult out large le front* (Roland v. 1217)
 between the two eyes much/very had large the brow
 “He had a large brow between his eyes”

It could be argued that (iii) does involve a small clause. Reasons of space prevent us from undertaking an in-depth comparison of Old French and Old Italian, which we must leave for now for future research.

- (9) a. ***molto** avea perduto **sangue**
 much had lost blood
 a'. ***avea molto** perduto **sangue**
 had much lost blood
 (cf. French *J'ai beaucoup perdu de sang*)
 b. *Il re **molto** ha grande dolore.
 The king very has big grief
 "The king was in very great pain"

The fact that *molto*_{EX} occurs only when adjectives and nominals are in predicative positions is an essential thing that must be captured by the analysis.

3. The movement path of *molto*_{EX}

In analyzing the path of *molto* when it extracts out of the small clause let us now consider its target position in the clause, which is rather easy to determine, since we have a rather precise idea of the left periphery in general, and of the left periphery in OI.

By and large, the vast majority of cases of *molto*_{EX} is located in a position immediately before the tensed verb (pre-T), i.e., in the CP layer, since OI is a verb second language.

- (10) *Certo molto ne sarebbe gran disinore se ...*
 Surely very of.it=would.be big dishonour if...
 "It would certainly be greatly dishonourable if ..."
 (VeV, cap. 58)

As the above example and those in (2a, b) and (4) attest, the only elements that can intervene between *molto* and the inflected verb are clitics, which can only occur in proclisis: no cases of enclisis, i.e., *molto*-V-cl are attested.

According to Benincà's (2006: 57) syntactic interpretation of the Tobler-Mussafia Law, proclisis signals that the Specifier of the Operator position located immediately above FinP is occupied (Benincà's (2006) FocusP). Since *molto* triggers proclisis, it must be located in Spec,Op. For this reason, we identify this position as OpP in the Focus Field (see also ModP in Rizzi 2004 and subsequent work for a dedicated projection for adverb modifiers, a position which hosts quantificational elements and probably also *wh*-items).

- (11) a. *Certo molto ne sarebbe gran disinore se ...*
 b. [_{Frame} [_{Force} [_{Topic} [_{Focus} [_{OpP} *molto* [_{FinP} [_{TP} ... [_{VP} [_{VP} ...] ...]]]]]]]]⁶

6. Out of the 275 relevant examples (see fn. 4), there are just 12 cases of the order XP-*molto*_{EX}-T, and 4 cases of *molto*_{EX}-XP-T. Though the numbers are very small, we can tentatively suggest that *molto*_{EX} targets a low projection in the CP.

In partial support of this claim we note the fact that CP *molto*_{EX} never co-occurs with the focus particle *si* that is claimed to occupy a low position in the CP layer (FocP in Benincà 2006: 17; OpP in Poletto 2014: 33). In other words, *molto*_{EX} competes with *si* for the same position in the CP (see Rossi & Poletto forthcoming).

The other possible position is between the inflected verb and the past participle.

- (12) a. *Questo Ruggieri di Loria era molto stato gran nemico de la Chiesa*
 This R. of L. was very been big enemy of the church
 “Ruggieri of Loria had been a huge enemy of the Church”
 (Paolino Pieri, p. 60)
- b. *L'opere della natura sono molto venute meno.*
 the works of=the nature are very come less
 “The works of nature have waned very much.” (Seneca, p. 91)

Cases of post-verbal *molto*_{EX} are far more limited (17 out of 275 total cases) than cases of CP-*molto*_{EX}, however, they are represented in the corpus. In cases like (12) *molto*_{EX} occupies the same position as adverbial bare *molto* (as other bare QPs like *tutto* and *niente*) as illustrated in (13), see Garzonio & Poletto (2012, 2017, 2018) and Poletto (2014) for *tutto* and *niente*.

- (13) [CP ... [TP ... [AdvP *molto* [vP [Topic [Focus [VP ...]] ...]]

This proposal is supported by the observation that bare quantifiers generally occur in front of the low left periphery containing Topics and Foci (as proposed by Belletti 2004: 18 for ModIt and Poletto 2014 for OI:56), and indeed there are examples that show that this is the case also for *molto*:

- (14) *Catuno de' detti partiti a pigliare è molto per Anibale reo, ...*
 Each of=the said parties to take.inf is very for Hannibal bad...
 “It would be very bad for Hannibal to take any of the parties ...” (FR, cap. 73)

Such a position, as we said, is the same targeted by bare adverbial *molto*, that is, the exact Manner Adverb position Cinque (1999: 11) proposes for ModIt *molto* and French *beaucoup*. More generally, we follow Cinque in assuming that bare QPs target dedicated projections in the IP domain of the clause.⁷

Once we have identified its target position, we can determine the exact mechanism that allows *molto*_{EX} out of a small clause but not out of a DP in OI and no extraction whatsoever in ModIt. The first difference between OI and ModIt has to

7. The difference between OI and modern Italian, where *molto* is located after the past participle, has to do with the fact that the past participle raises less in OI (as it does in ModIt, see Cinque 1999: 45 ff.).

do with the fact that *molto* could modify a prenominal adjective in OI but cannot in the modern language (Poletto 2014: 82ff):

- (15) a. *Allora il garzone, ..., pensò una molto bella buscia*
 then the boy ..., thought a very nice lie ...
 (FSI, p. 137)
- b. *Cornacchie sono di molto grande vita*
 crows are of very long life
 “Crows have a long lifespan”
 (Tesoro, b175)

This is most probably related to the fact that restrictive adjectives could be prenominal in OI but cannot nowadays:

- (16) a. *di vendichare la ricevuta onta*
 of avenge.INF the received shame
 “to avenge the offence received”
 (Distruzione di Troia, 164)
- b. *nel mio pecto dalla sinistra parte*
 in=the my chest from=the left part
 “in the left part of my chest”
 (VN 71, 3)

Furthermore, scrambling is possible also within the adjectival phrase, as shown by the following examples where the argument of the adjective has raised past the adjective itself:

- (17) a. *ch'è tanto di valor pieno*
 that is much of valor full
 “who is really full of valor”
 (Cavalcanti, Rime 561)
- b. *Loda colui, che di poco è ricco, e ...*
 prize him-there, that of little is rich and
 “Prize the one, who is rich with few things, ...”
 (Seneca 296)

This is also true of the adjectival head, which as we noticed above in (7) (here repeated as (18a)) can strand *molto* (or their internal argument), but also its complement as shown by (18b):

- (18) a. *e di gentile aspetto molto*
 and of kind appearance very
 (VN, cap. 8, par. 1, v. 11)
- b. *e ciò non è propria natura di cavallo*
 and this not is own nature of horse
 “and this is not the nature typical of a horse”
 (Nov., p. 128, r. 67)

This provides us with a way to understand why *molto* can only be extracted out of a predicative nominal expression, but not out of a regular definite DP. The

remnant AdjP containing only *molto* can be moved to reach either the adverbial position of bare quantifiers in the low IP area or to the operator position in the CP left periphery only if there is no phase boundary that blocks it, i.e., when the nominal expression is not a full-fledged DP. When DP is projected, movement is blocked by the phase boundary. However, one could in principle use SpecD as an escape hatch for *molto*, as we regularly assume to be the case for the vP phase. One might wonder why DPs are phases without a visible edge that can be used as the escape hatch for further movements into the clausal spine. We surmise that this must be related to the fact that there is a definite article in the DP projection that blocks movements of other elements rendering the DP phase opaque.⁸ This means that OI is similar to ModIt in not allowing movement of an adjective out of the DP phase, i.e., in having opaque DPs. However, it can still move it if the phase boundary is not present, i.e., in predicative constructions. Now, the next problem is why ModIt does not allow for *molto*_{EX} in case of predicative structures. The solution to this problem is very simple: ModIt does not have access to the CP left periphery in the same way OI did, the operator position above FinP only contains *wh*-items, and contrastive/corrective Focus, but no other elements. However, even those modern languages where modifiers can be found in the CP left periphery do not display *molto*_{EX}. The reasons for this block is that two conditions must be met for *molto*_{EX} to be grammatical: (a) the accessibility of the CP left periphery and (b) the accessibility of the high functional spine internal to the nominal. The second condition is not met even in languages like modern Sardinian and Sicilian. In other words, the problem has to do with the usual culprit, i.e., the loss of the V2 property intended in its widest interpretation, i.e., the accessibility of the left periphery of each phase (CP, vP; DP) not only by XPs internal to the phase but also to the lexical head/project (T, v, N).

8. There is actually just one case of extraction out of the DP that is possible in Italian, namely extraction of a *wh*-item which is the complement of the N as shown in (i)

- (i) *Di quale libro conosci l'autore?*
 Of which book know the author?
 "Of which book do you know the author?"

These cases are marginal, but still possible. This means that the opaqueness of the DP must be modularized with respect to the element that is extracted. Quantifiers cannot move out of the phase boundary, while *wh*-items can. As to why this is so, one can imagine a scenario in which the phase boundary is permeable to some features, but not others. We leave this point for future research.

4. Concluding remarks

In this chapter, we have shown that a very specific syntactic property, like scrambling of an adjectival modifier like *molto* in OI can ultimately be traced back to a better-known property, i.e., the one that makes OI a V2 language. Its loss in modern Italian can as well be traced back to the loss of verb second. The restriction of *molto*_{EX} to cases of predicative constructions is accounted for by assuming that the OI DP already had lost the properties that allowed for scrambling of a quantificational element like *molto* out of it, so that scrambling is possible only when the DP phase is missing, i.e., in predicative constructions. This loss is probably related to the development of the definite article, which changed the properties of the D projection. More generally, we have shown that scrambling within the nominal phase was indeed possible, since it could target an adjective but strand the complement of the adjective.

Acknowledgements

We would like to thank Jacopo Garzonio, Guido Mensching and Francesco Pinzin for relevant discussion, and the audience at Going Romance 2018, Utrecht for comments and suggestions.

Funding

This paper is part of the DFG-funded Research Project *Quantifikation im Altitalienischen* (PO 1642/2-2)). For the purposes of the Italian Academy, Silvia Rossi is responsible for §§ 2 and 3 and Cecilia Poletto for §§ 1 and 4.

References

- Abels, Klaus. 2003. Successive cyclicity, anti-locality, and adposition stranding. Ph.D. Dissertation: University of Connecticut.
- Baunaz, Lena. 2011. *The grammar of French quantification*. Dordrecht: Springer.
<https://doi.org/10.1007/978-94-007-0621-7>
- Belletti, Adriana. 2004. Aspects of the low IP area. In Luigi Rizzi (ed.), *The structure of CP and IP: The cartography of syntactic structures*, Volume 2, 6–51. Oxford: Oxford University Press.
- Benincà, Paola. 2006. A detailed map of the left periphery of Medieval Romance. In Raffaella Zanuttini, Héctor Campos, Elena Herburger & Paul Portner (eds.), *Negation, tense and clausal architecture: Cross-linguistic investigations*, 53–86. Washington, DC: Georgetown University Press.

- Camus Bergareche, Bruno & Manuel Pérez Saldanya. 2011. From adjectives to quantifiers: Evidence from Old and Modern Catalan. *Catalan Journal of Linguistics* 10. 221–245.
<https://doi.org/10.5565/rev/catjl.62>
- Carlier, Anne. 2011. From *multum* to *beaucoup*: Between adverb and nominal determiner. In Lucia Tovena (ed.), *French determiners in and across time*, 55–87. London: College Publications.
- Chomsky, Noam. 2005. On Phases. In Robert Freidin, Carlos P. Otero & Maria Luisa Zubizarreta (eds.), *Foundational issues in linguistic theory*, 133–166. Cambridge, MA: MIT Press.
- Corver, Norbert. 2017. Subextraction. In Martin Everaert & Henk van Riemsdijk (eds.), *The Wiley-Blackwell companion to syntax*, 4139–4189.
<https://doi.org/10.1002/9781118358733.wbsyncom054>
- De Clerq, Karen. 2017. The internal syntax of Q-words. *Linguistics in the Netherlands* 34(1). 31–46. <https://doi.org/10.1075/avt.34.03dec>
- Garzonio, Jacopo & Cecilia Poletto. 2012. On *niente*: Optional negative concord in Old Italian. *Linguistische Berichte* 230. 131–153.
- Garzonio, Jacopo & Cecilia Poletto. 2017. How bare are bare quantifiers? Some notes from diachronic and synchronic variation in Italian. *Linguistic Variation* 17(1). 44–67.
<https://doi.org/10.1075/lv.17.1.03gar>
- Garzonio, Jacopo & Cecilia Poletto. 2018. The distribution of quantifiers in Old and Modern Italian: Everything or nothing. In Ana Maria Martins & Adriana Cardoso (eds.), *Word order change*, 221–239. Oxford: Oxford University Press.
- Giorgi, Alessandra. 2010. La struttura del sintagma nominale. In Giampaolo Salvi & Lorenzo Renzi (eds.), *Grammatica dell'italiano antico*, 275–294. Bologna: il Mulino.
- Giusti, Giuliana. 2010a. Le espressioni di quantità. In Giampaolo Salvi & Lorenzo Renzi (eds.), *Grammatica dell'italiano antico*, 377–400. Bologna: il Mulino.
- Giusti, Giuliana. 2010b. Il sintagma aggettivale. In Giampaolo Salvi & Lorenzo Renzi (eds.), *Grammatica dell'italiano antico*, 593–616. Bologna: il Mulino.
- Leu, Thomas. 2008. 'What for' internally. *Syntax* 11. 1–25.
<https://doi.org/10.1111/j.1467-9612.2008.00113.x>
- Longobardi, Giuseppe. 1994. Reference and proper names: A theory of N-movement in syntax and Logical Form. *Linguistic Inquiry* 25(4). 609–665.
- Obenauer, Hans-Georg. 1994. *Aspects de la syntaxe A-barre*. Thèse de doctorat d'État, Université Paris VIII.
- Poletto, Cecilia. 2014. *Word order in Old Italian*. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199660247.001.0001>
- Poletto, Cecilia. 2018. L'etimologia come finestra sulla sintassi, evidenza empirica dai dialetti italiani. In Martin-Dietrich Glessgen, Johannes Kabatek & Harald Völker (eds.), *Repenser la variation linguistique. Actes du Colloque DIA IV à Zurich (12–14 sept. 2016)*, 107–121. Strasbourg: Éditions de linguistique et de philologie.
- Rizzi, Luigi. 2004. Locality and left periphery. In Adriana Belletti (ed.), *Structures and beyond: The cartography of syntactic structures*, Volume 3, 223–251. New York: Oxford University Press.
- Rohlf, Gerhard. 1969. *Grammatica storica della lingua italiana e dei suoi dialetti. Sintassi e formazione delle parole*. Turin: Einaudi.
- Ross, John Robert. 1967. Constraints on variables in syntax. Ph.D. Dissertation: Massachusetts Institute of Technology.

- Rossi, Silvia & Cecilia Poletto. Forthcoming. Bare quantifiers and verb second: The view from Old Italian. In Nicholas Catasso, Marco Coniglio & Chiara De Bastiani (eds.), *Language change at the interfaces. Interasential and intersentential phenomena*. Amsterdam: John Benjamins..
- Salvi, Giampaolo & Lorenzo Renzi (eds.). 2010. *Grammatica dell'italiano antico*. Bologna: il Mulino.
- Talić, Aida. 2017. Adverbial left-branch extraction and the structure of AP in Slavic. In Yohei Oseki, Masha Esipova & Stephanie Harves (eds.), *Annual workshop on formal approaches to Slavic linguistics (FASL), The NYU Meeting 2015*, 399–418. Ann Arbor, MI: Michigan Slavic Publications.

Sources

- DCompagni* = Dino Compagni, *Cronica delle cose occorrenti ne' tempi suoi (Dino Compagni e la sua Cronica*, edited by Isidoro Del Lungo, Volume 3, Florence, Le Monnier, 1887).
- De Amore* = Anonymous, *De Amore di Andrea Cappellano volgarizzato* (a cura di Graziano Ruffini, Milan, Guanda, 1980)
- Distruzione di Troia* = Anonymous, *Da un libro della distruzione di Troia (Testi fiorentini del Duecento e dei primi del Trecento*, edited by Alfredo Schiaffini, Florence, Sansoni, 1926, pp. 151–84)
- FR* = Bono Giamboni, *Fiore di rettorica (redazione beta)* (Bono Giamboni, *Fiore di rettorica*, edited by Gian Battista Speroni, Pavia, Dipartimento di Scienza della Letteratura e dell'Arte medioevale e moderna, 1994, pp. 3–107)
- FSI* = Anonymous, *Fiori e vita di filosafi e d'altri savi e d'imperadori*, (edited by Alfonso D'Agostino, Florence, La Nuova Italia, 1979)
- Nov.* = Anonymous, *Il Novellino*, 13th cent. (edited by Guido Favati, Genoa, Bozzi, 1970).
- Paolino Pieri* = Paolino Pieri, *Cronica* (edited by Anton Filippo Adami, Rome, Monaldini, 1755).
- Rettorica* = Brunetto Latini, *La Rettorica* (edited by Francesco Maggini, Florence, Le Monnier, 1968).
- Rime Cavalcanti* = Guido Cavalcanti, *Rime (Poeti del Duecento*, edited by Gianfranco Contini, Milano-Napoli, Ricciardi, 1960, t. II, pp. 491–558, 561, 563–64, 566–67).
- Seneca* = Anonymous, *Pistole di Seneca volgarizzate (Volgarizzamento delle Pistole di Seneca e del Trattato della Provvidenza di Dio*, edited by Giovanni Bottari, Florence, Tartini e Franchi, 1717, pp. 1–418)
- Tesoro* = *Il Tesoro di Brunetto Latini volgarizzato da Bono Giamboni* (raffrontato col testo autentico francese edito da P. Chabaille, emendato con mss. ed illustrato da Luigi Gaiter, Bologna, Presso Gaetano Romagnoli, Volume 4, 1878–83)
- VeV* = Bono Giamboni, *Il Libro de' Vizi e delle Virtudi (Il Libro de' Vizi e delle Virtudi e Il Trattato di Virtù e Vizi*, edited by Cesare Segre, Turin, Einaudi, 1968, pp. 3–120)
- VN* = Dante Alighieri, *La Vita Nuova* (edited by Michele Barbi, Florence, Bemporad, 1932)

On adverbial perfect participial clauses in Portuguese varieties and British English

Purificação Silvano¹, António Leal¹ and João Cordeiro^{2,3}

¹Universidade do Porto and CLUP/ ²Universidade de Beira Interior /

³INESC TEC – Porto

The purpose of this paper is twofold: (i) to discuss the validity of Lobo's proposal (2003) of distinguishing two types of adverbial perfect participial clauses (APC) in European Portuguese; and (ii) to ascertain the key factors behind their temporal interpretation. To achieve these aims, we compare and contrast patterns displayed by APC in European, Brazilian, Mozambican and Angolan Portuguese and British English across a corpus built from newspapers. Our research reveals that the data do not reflect the bipartite division argued for by Lobo and that, for Portuguese varieties, the position of APC in the sentence and the combinations of some aspectual classes are important to infer temporal relations, whereas, for British English, the anterior orientation/perfect aspect of the perfect participle imposes for most cases an anteriority temporal relation, surpassing the influence of any other factor.

Keywords: adverbial perfect participial clauses, corpus-based study, Portuguese varieties, British English, typology, temporal interpretation, aspectual classes, ordering, tense

1. The problem

Adverbial perfect participial clauses (APC), which have an auxiliary verb in the *-ing* form followed by the past participle of the main verb,¹ have been the object of syntactic and semantic research as part of larger studies both for Portuguese and

-
1. The structure under scrutiny is labeled differently in the two languages at hand: while in Portuguese they are called adverbial gerundive clauses, in English they are traditionally known as adverbial perfect participial clauses. For English, there has been some discussion as to the two-fold distinction of *-ing* clauses based on the presence of a gerund or a participle (cf. De Smet 2010). The shortcomings of both terms provide arguments for Haspelmath (1995) suggesting the term 'converb'. In order to avoid misunderstandings, we have deemed it best to opt for the labels 'adverbial perfect participial clauses' and 'perfect participle' when referring to this structure in both languages.

English (for English, Quirk et al. 1985; Stump 1985; Kortmann 1995; König 1995; Sæbø 2011, a.o.; for Portuguese, Leal 2002; Lobo 2003; Mória & Viotti 2004a; Cunha, Leal & Silvano 2015; Leão 2018; López 2019, a.o.), including typological proposals.

However, these proposals are not without problems. In particular, when one looks into the prevailing proposal to classify European Portuguese adverbial perfect participial clauses from Lobo (2003, 2006, 2013), it does not fully account for the data. In a nutshell, Lobo (2003, 2006, 2013) puts forward the following division for European Portuguese APC: (i) peripheral (or sentence) clauses (1a), typically occupying an initial position (or a final position, if moved) with pause between clauses and conveying a temporal value of anteriority; (ii) and coordinate (posteriority) clauses (1b), occurring only in final position and exhibiting a temporal value of posteriority.

- (1) a. *Tendo chegado atrasado, o Zé só encontrou lugar na última fila.*
 Having-GER² arrived-PTCP late, the Zé only found seat in-the last row
 “Having arrived late, Zé only found a seat in the last row.”³
- b. *Os bandidos escaparam à polícia, só tendo sido identificados dois dias depois.*
 The bandits escaped to-the police, only having-GER been-PTCP identified two days later
 “The bandits escaped the police, having only been identified two days later.”
 (Lobo 2003: 249)

The problem with this proposal is that there are data, like (2)–(4), that seem to contradict this division. In examples (a), the same APC occurs in the final position, that is, after the main clause, with different temporal readings: anteriority (2a) and posteriority (3a, 4a). However, when the order of the clauses is changed, like in examples (b), different results are obtained: in (2b), the APC continues to be interpreted as anterior to the main clause; in (3b), the temporal reading changes with the order of the clauses, the escape occurring after the capture; in (4b), placing the APC in an initial position produces an ungrammatical, or at least strange, sentence, because it is odd to interpret the situation of the subordinate clause as occurring before the situation of the main clause.

-
2. ‘GER’ stands for ‘gerund’ (or present participle in English terminology).
 3. Our translation.

- (2) a. *Os ladrões passaram uma noite na cadeia, tendo sido capturados na passada segunda-feira.*
 The thieves spent a night in-the jail, having-GER been-PTCP captured in-the last Monday
 “The thieves spent a night in jail, having been captured last Monday.”
- b. *Tendo sido capturados na passada segunda-feira, os ladrões passaram uma noite na cadeia.*
 Having-GER been-PTCP captured in-the last Monday, the thieves spent a night in-the jail
 “Having been captured last Monday, the thieves spent a night in jail.”
- (3) a. *Os ladrões fugiram da prisão por uma janela, tendo sido capturados na passada segunda-feira.*
 The thieves fled of-the prison through a window, having-GER been-PTCP captured in the last Monday
 “The thieves fled the prison through a window, having been captured last Monday.”
- b. *Tendo sido capturados na passada segunda-feira, os ladrões fugiram da prisão por uma janela.*
 Having-GER been-PTCP captured in-the last Monday the thieves fled of-the prison through a window
 “Having been captured last Monday, the thieves fled the prison through a window.”
- (4) a. *Os ladrões andaram fugidos durante dois dias, tendo sido capturados na passada segunda-feira.*
 The thieves were escaped for two days having-GER been-PTCP captured in-the last Monday
 “The thieves were on the run for two days, having been captured last Monday.”
- b. **/???Tendo sido capturados na passada segunda-feira, os ladrões andaram fugidos durante dois dias.*
 the thieves were escaped for two days
 “Having been captured last Monday, the thieves were on the run for two days.”

These examples are problematic for Lobo’s proposal (2003, 2006, 2013), since, even if one considers that the same APC in (2) is a peripheral clause but in (4) is a coordinate clause, her classification cannot explain why in (3) the APC carries a different temporal interpretation. So, should we consider APC in (3) as a peripheral APC placed at the end of the sentence or as coordinate APC? Both alternatives raise issues with Lobo’s analysis, because if we assume that it is a peripheral APC, then

it exhibits a posteriority meaning not predicted by Lobo, and, if we accept it as a coordinate APC, then it could only occupy final position.

Therefore, one of the aims of this study was to determine if the problems that we have just highlighted regarding Lobo's characterization of European Portuguese APC are borne out by a corpus study and are statistically significant. Furthermore, we also wanted to verify if the same issues occur in APC not only in other varieties of Portuguese (PT), but also in English.

The choice of PT varieties in this study took into consideration the fact that European Portuguese (EP) and Brazilian Portuguese (BP) have stabilized norms, unlike Angolan Portuguese (AP) and Mozambican Portuguese (MP). Despite the fact that AP and MP do not have a stabilized norm, it is relevant to ascertain to what degree these varieties are similar (or not) to BP or EP, since they are probably influenced by other African languages (namely Bantu languages), which are typologically different from PT (cf. Gonçalves 2010, 2013; Carvalho & Lucchesi 2016).

The choice of British English (BE) is justified by the fact that, although this language has a similar construction to PT, it is normally said to be marked for perfect aspect, that is, with an anterior orientation, and a temporal interpretation of succession of situations (Kortmann 1995: 221). Thus, the possibility of other temporal relations is not described for BE, contrary to PT. Additionally, Stump (1985) argues for a semantic analysis of the perfect, which accounts for the temporal interpretation of the perfect in free adjuncts, proposing that the function of the perfect is "to locate an event within a perfect interval" (Stump 1985: 247):

a perfect sentence ϕ without a time adverb is true at interval i iff the tenseless form ψ of ϕ is true sometime during a perfect interval relative to i ; a perfect sentence ϕ with a time adverb α is true at i iff there is some perfect interval which is in the set of intervals determined by α and during which ψ is at some time true.

(Stump 1985: 232)

In fact, for PT, the perfect participle is usually described as perfective as well, indicating that a situation is concluded before the one expressed by the main clause (Cunha & Cintra 1984: 487) and anterior to a reference point, behaving, thus, similarly to other perfective tenses, namely the *infinitivo composto* ("perfect infinitive"), which combines the auxiliary verb *ter* ("to have") in the infinitive and the past participle of the main verb, and the *particípio passado* ("past participle") in absolute constructions, i.e., adverbial clauses that combine this verb form together with the verb's internal argument lexically fulfilled (Oliveira 2013: 551).⁴

4. See also Peres (1994), who also argues that the past participle in EP is always perfective.

Challenging this view, Leal (2002, 2011) claims that the temporal values of APC in EP can be explained using Kamp & Reyle's (1993) two-dimensional theory of tense. He assumes that they are usually temporally located in the following way: the 'Temporal Perspective' is the time interval of the main clause; the 'TENSE' is always 'past'. This means that (i) the relationship between the Temporal Perspective Point (TPpt) of the APC and the utterance time is expressed by the tense of the main clause, and (ii) the time interval that locates the situation described by the APC is always anterior to the TPpt. So, both in (5) and (6) the situation denoted by the APC is located in a time interval anterior to the one that locates the situation in the main clause ('Tense' is always 'past'). However, due to the tense of the main clause (*Pretérito Perfeito* ("Simple Past") in (6) and *Futuro* ("Future") in (7)) the 'Temporal Perspective' is different ([+PAST] in (6) and [-PAST] in (7)).

- (5) *Tendo destruído a cidade, o exército inimigo*
 Having-GER destroyed-PTCP the city, the army enemy
avançou 10 km.
 move-PST 10 km
 "Having destroyed the city, the enemy army moved forward 10 km."
- (6) *Tendo destruído a cidade, o exército inimigo*
 Having-GER destroyed-PTCP the city, the army enemy
avançará 10 km.
 move-FUT 10 km
 "Having destroyed the city, the enemy army will move forward 10 km."

However, when the APC are placed after the main clause with the verb in the *Pretérito Perfeito* (Simple Past), they can exhibit, in addition to the anteriority reading, also a posteriority one (see the compatibility with temporal adverbials *antes/depois* ("before"/"after") in (7)).

- (7) *O exército inimigo avançou 10 km, tendo destruído*
 The army enemy move-PST 10 km, having-GER destroyed-PTCP
{antes/depois} a cidade.
 {before/after} the city
 "The enemy army moved forward 10 km, having destroyed {before / after} the city."

Leal (2011) assumes that the relevant TPpt to the posteriority reading in (7) is the utterance time based on examples like (8). In this case, the tense in the main clause locates the situation in a time interval after the utterance time. Under these circumstances, the utterance time is not available as the TPpt of the APC, only the time interval locating the main clause, and the posteriority reading is blocked.

(8) *O exército inimigo avançará 10 km, tendo destruído*

The army enemy move-FUT 10 km, having-GER destroyed-PTCP

{antes/*depois} a cidade.

{before/after} the city

“The enemy army will move forward 10 km, having destroyed {before/*after} the city.”

So, according to Leal (2002, 2011), the temporal relations established by APC cannot be said to rely solely on the temporal value of the perfect participle, depending on an array of factors such as the order of clauses and tense of the main clause.⁵ In fact, for English, also König (1995: 73–85) maintains that the interpretation of perfect participial constructions results from “the interaction between a basic vague meaning of the converb and a wide variety of the syntactic, semantic and contextual factors”, namely grammatical properties (order of the clauses); semantic factors (connectors, tense, aspect, mood, and modality); and pragmatic factors (world knowledge; general background assumptions and contextual information; general principles of language).

Bearing in mind what we have just presented, one cannot help but wonder not only what the exact contribution of the perfect participle to the temporal relations of APC in PT and BE is, but also what other factors are relevant to the temporal interpretation of these clauses and can be assumed as criteria to define a typology of APC.

Therefore, the second aim of our study, intertwined with the first, was to establish the relevant factors to the temporal interpretation of APC. In this study, we confine ourselves to the following factors: position (initial, medial, final), tense of the main clause (MC) and aspectual types of the MC and subordinate clause (SC).

Although, as mentioned before, there are studies on the semantics of APC, mainly for EP and BE, there is not, as far as we know, a comprehensive comparative analysis between not only different PT varieties, but also between these and BE. Such analysis is of utmost importance to better understand the semantic behavior of these structures in both languages so that a more grounded and complete description (and, eventually, formal classification) can be put forward, since there are questions which cannot be explained or are not predicted by current proposals.

This paper is organized as follows: § 2 describes the data and the method behind the constitution of the corpus and the analysis; § 3 presents and discusses the findings of our research; and § 4 summarizes the main results.

5. Concerning EP, see also Lobo (2006), among others.

2. Methodology

2.1 The data and method

The corpus was built automatically by a web crawler (Risvik & Michelsen 2002; Najork 2009; Olston & Najork 2010) that was specially designed and tuned for scanning specific online newspapers from different languages/varieties (see Table 1), and for selecting sentences satisfying linguistic patterns, that is, APC. Basically, a crawler is a piece of software that systematically selects data (mainly text) from the Web. In our case, from a given newspaper, the crawler selects sentences satisfying the desired syntactic criteria. Besides that, for each page the crawling process collects all internal⁶ links, which afterwards are followed, and so recursively processed, until a predefined depth level is reached. Our crawler was implemented in Java and uses minimum linguistic resources for newspaper's sentence selection, namely, a statistically induced part-of-speech tagger (Klein & Manning 2003), using models for English and Portuguese, and a rule based lemmatizer (Muniz, Nunes & Laporte 2005).

Table 1. Newspapers used in the data gathering

Country	Newspapers	Number N
Portugal	expresso.sapo.pt; observador.pt; dn.pt; ionline.pt; publico.pt	5
Brazil	agazetadoacre.com; diariobahia.com.br; diariodoamazonas.com.br; g1.globo.com; jmonline.com.br; jornaldosol.com.br; jornalgazetadooeste.com.br; correio24horas.com.br; correiobraziliense.com.br; em.com.br; estadao.com.br; jornalestadodegoias.com.br; oriobranco.net; unicamp.br	14
Angola	angola24horas.com; apostoladoangola.org; jornaldeangola.sapo.ao; jornal8.net; tpa.sapo.ao; angop.ao; correioangolense.com; novojournal.co.ao	8
Mozambique	verdade.co.mz; noticias.mmo.co.mz; opais.sapo.mz	3
UK	metro.co.uk; telegraph.co.uk; theguardian.com; express.co.uk; thetimes.co.uk	5

Five geographical regions were considered: for PT, Portugal, Brazil, Angola, and Mozambique; for BE, the United Kingdom. A total of 4222 and 2635 relevant sentences were extracted during November 2018 (see Table 2).⁷

6. Links pointing to other pages from the same site.

7. In certain cases, such as with BP, it was necessary to resort to more newspapers because there was less text in each newspaper than in other varieties/languages.

Table 2. Web crawler results

Country	Extracted	Sentences	Words	Extracted/ Sentences (%)	Words/ Sentences
Portugal	2 587	1 396 535	39 450 278	0.1852%	28,25
Brazil	235	631 809	15 881 679	0.0372%	24,14
Angola	857	611 855	24 247 326	0.1401%	32,22
Mozambique	543	171 693	5 114 459	0.3163%	29,79
UK	2 635	1 732 969	38 597 456	0.1522%	22,27

The next step was the selection of 200 random samples of sentences with APC for each variety/language.⁸ The corpus was analyzed with respect to the following parameters: temporal relation (anteriority, simultaneity, posteriority); position (initial, medial, final); tense of the main clause (MC); aspectual types of the MC and subordinate clause (SC):⁹ State (St), Culminated Process (CP), Process (Pro), Culmination (Culm), Point (Pon).

3. Discussion of findings

This section presents a quantitative analysis of APC found in the corpus and the discussion of the results, focusing on the relationship between temporal readings of APC and their position, the tense of the MC and aspectual classes of both subordinate and main clauses.

3.1 Temporal readings and position

The preferential position in all varieties/languages is the final one, although there are some differences concerning BE data, where the distribution between the three possibilities is more balanced (see Figure 1). The statistical χ^2 tests¹⁰ revealed significant differences between all language/variety pairs, with the exception of EP-AP, EP-MP, and AP-MP.

8. The exception was BP, because after deleting the non-relevant data, the final output was 191 sentences.

9. Here we use Moens' (1987) ontology. In a nutshell, situations can be described as statives or eventives. In the latter case, that means that situations are constituted by one or more phases, which reflects on the aspectual properties of predications. So, Processes are durative and atelic; Culminated Processes are durative and telic; Culminations are non-durative and telic; finally, Points are non-durative and atelic.

10. In the empirical measurements reported in this article, a Chi-Squared test was applied to each pair, using a significance level of 0.05, meaning a rejection of the null hypothesis (H_0) if we have 'p-value' < 0.05.

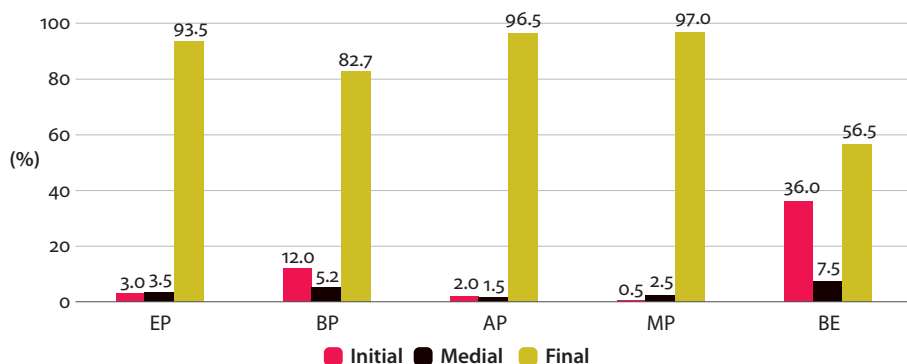


Figure 1. Position of APC

Examples (9)–(11), from the BE corpus, illustrate these different positions.

- (9) *Having initially denied responsibility*, Prause then told detectives he became angry after Coleman made a jokey comment about him needing to grow up and responded by attacking her with a hammer he was holding.
- (10) But to understand the forces that have brought the company to its knees, we can begin in that exciting period of the mid-19th century when newspapers – *having been freed from the so-called taxes on knowledge* – were the sole means of mass communication.
- (11) Vichai, 60, was a hugely popular figure at Leicester, *having bought the club for £ 39m in 2010*.

Regarding temporal relations, EP, AP and MP show a preference for posteriority readings, whereas BP and BE prefer anteriority readings (clearer preference in BE) (see Figure 2). All the differences between the languages/varieties are significant from a statistical point of view.

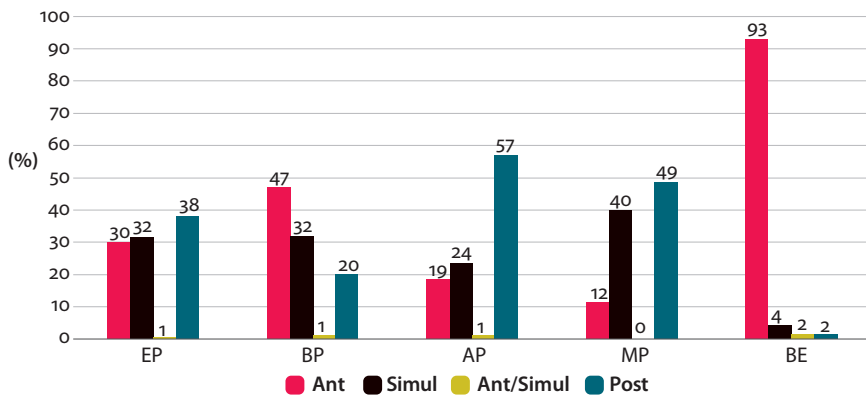


Figure 2. Temporal relations of APC

Examples (12)–(17) illustrate these readings in PT and BE languages.

- (12) *Foi duas vezes vereador de Salvador, tendo sido eleito pela primeira vez em 1988.*
 Was two times councilor of Salvador, having-GER been-PTCP elected
for-the first time in 1988.
 “He was twice councilor of Salvador, having been elected for the first time in 1988.” (BP, anteriority)
- (13) In another meeting, Mark Turnbull, managing director of CA Political Global, described how, *having obtained damaging material on opponents*, the company can discreetly promote it on to the internet. (BE, anteriority)
- (14) *João Lourenço percorreu as áreas de marcenaria, soldadura ... e mecânica auto, tendo conversado com os formandos e formadores.*
 João Lourenço toured the areas of carpentry, welding, ... and mechanics auto, having-GER talked-PTCP with the trainees and trainers
 “João Lourenço toured the areas of carpentry, welding ... and auto mechanics, having talked with the trainees and trainers.” (AP, simultaneity)¹¹
- (15) Before joining GNM she has had a long and varied career in media and marketing *having spent seven years at Google and YouTube ...* (BE, simultaneity)
- (16) *A paciente ... seria submetida a uma ecografia que determinou a ocorrência de uma gravidez extra-uterina, tendo sido extraído à cesariana um feto vivo.*
 The patient ... be submitted to an ultrasound that determined the occurrence of a pregnancy extrauterine, having-GER been-PTCP extracted by-the cesarean section a fetus live

11. When we identify a simultaneous reading in the corpus, this means that the two situations are located in time intervals that overlap partially or totally. See, for instance, Example (14), from the Angolan corpus. In this case, the main clause corresponds to a Culminated Process, whereas the APC corresponds to a Process. Or, using Vendler’s terms, Accomplishment and Activity, respectively. The usual tests were used to identify the aspectual classes. For instance, one can say *João Lourenço percorreu as áreas ... in × time*, but one can only say *tendo conversado com os formandos ... for × time* (cf. Dowty 1979, and, for EP, Cunha 2004, or Leal 2009). Back to simultaneity reading, we can explain it by resorting to Moens’ (1987) notion of aspectual nucleus: the preparatory processes of the main clause and APC occur simultaneously (so, it is not the case that the simultaneous reading arises because the preparatory process of the APC occurs during the consequent state of the main clause). Therefore, we can say that João Lourenço talked with the trainees and trainers during the tour to the abovementioned areas, and not after the tour. Finally, note that, for English, these simultaneity readings were validated by two native speakers.

“The patient ... would be submitted to an ultrasound that determined the occurrence of an extrauterine pregnancy, a live fetus having been extracted by cesarean section.”
(MP, posteriority)

- (17) Salaam served more than five years in prison for a crime that he didn't commit, *having been released in 2002 along with the others of the Central Park Five*.
(BE, posteriority)

Looking into the combination of the position of the SC and the temporal relation (see Table 3), we observe that, when the SC occupies the initial or medial position, the most common temporal relation is of anteriority in all varieties/languages. However, when the SC is placed in final position, differences arise: in BP and BE, the preferential temporal relation is of anteriority (although the differences regarding the other temporal readings are more notorious in BE than in BP), while in EP, AP and MP it is of posteriority. A χ^2 test shows that all the differences are statistically significant, except for the pair AP-MP.

Table 3. Results of the pair position/temporal relation

Position of SC/Temporal relation	EP	BP	AP	MP	BE
Initial-Ant	6	22	2	1	70
Initial-Simul	0	1	1	0	0
Initial-Post	0	0	1	0	0
Medial-Ant	6	8	3	4	15
Medial-Simul	1	2	0	0	0
Medial-Post	0	0	0	1	0
Final-Ant	48	60	32	18	101
Final-Simul	62	58	46	80	8
Final-Post	76	38	113	96	3

According to these data, the position of the clause seems to correlate with temporal relation in EP, AP, and MP, since the temporal ordering of eventualities tends to follow the clauses ordering in the sentence. This relation is not so obvious in BP, because most of the APC in final position are interpreted as temporally anterior to the situation of the MC. Clearly it is not the case in BE, where the anterior relation arises irrespective of the position in the sentence.

The analysis we have just presented allows us to make some noteworthy observations concerning the two problems posed in § 1: (i) Lobo's classification of EP APC (2003) and its validity in relation to BE APC; (ii) and the pertinence of the temporal/aspectual nature of the perfect participle, position (one of the criteria to define the type of APC within Lobo's proposal), tense and aspectual classes in determining temporal interpretation of the APC.

Regarding Lobo's proposal, the PT data provide some counterevidence. Most APC occur in final position; if the main temporal interpretation was posteriority, they should be coordinate clauses; however, that only happens in most cases in AP and MP. On the contrary, in EP and BP, most of the temporal readings correspond to anteriority and simultaneity readings. This is odd, because, if we were to follow Lobo's analysis, this would mean that, in EP and BP, most of the APC do not occur in their basic position, but are moved to a position after the main clause.

What we would expect from Lobo's proposal was a distribution like the one we find in BE, as there is a more balanced distribution of positions (between initial and final position). However, on the other hand, it seems that extending the bipartite analysis of APC argued by Lobo for PT to BE would not be very productive, because anteriority seems to prevail regardless of the position, posteriority readings being scarce in this language.

As to the factors underlying the readings of APC, these data show that, in BE, contrary to PT, position is not a relevant factor, since there is not a relation between position and temporal reading. In this case, the perfective value of the perfect participle seems to override the factor position.

3.2 Temporal readings and tense of the main clause

As to the role of tenses of the MC in the temporal interpretation,¹² the following conclusions can be drawn: in EP and the African varieties, the most frequent is PP with the reading of posteriority of the APC, followed by PP – Simul. BP behaves again differently, being the most recurrent examples of Pres-Ind – Ant, followed by PP – Simul (see Table 4). The χ^2 test revealed significant differences for all varieties' combinations.

Table 4. Results of the pair tense MC/temporal relation in PT

Tenses/Temporal relation	EP	BP	AP	MP
PP – Post	66	30	109	89
PIMP-Ind – Simul	6	7	7	6
PP – Ant	20	35	15	14
PP – Simul	43	42	30	59
PPC-Ind – Simul	6	3	3	2
Pres-Ind – Ant	29	44	16	8
Pres-Ind – Simul	7	4	2	10

12. The tenses are represented by the following initials: for PT, *Pretérito Perfeito*- PP, *Presente do Indicativo* – Pres-Ind, *Pretérito Imperfeito* – PIMPInd, *Pretérito Perfeito Composto do Indicativo* – PPC-Ind, and, for BE, Simple Past with a perfective value – Pst, 'Present' – Pres, 'Present Perfect' – PresP.

Regarding BE, most cases correspond to anteriority readings, both with Pst and Pres. In comparison to PT varieties, BE gives distinct preference to the anteriority temporal relation with Pst, whereas PT favors a posteriority relation with a partially equivalent tense (see Table 5).

Table 5. Results of the pair tense of the MC/temporal relation in BE

Tenses/Temporal relation	BE
Pst-Ant	82
Pres-Ant	75
PresP-Ant	14
PresP-Simul	5

A closer look at the data with PP in PT, a tense that does not perform any aspectual change in the predication, reveals that PP can occur with all temporal relations, especially when the APC occurs after the main clause (see Table 6).

Table 6. Results of the combination tense of the MC with PP/temporal relation/position of SC in PT

Temporal relation	Position of SC	PP			
		EP	BP	AP	MP
Ant	Initial	2	11	2	1
	Medial	0	6	2	4
	Final	18	18	11	9
Simul	Initial	0	1	1	0
	Medial	0	1	0	0
	Final	43	40	29	59
Post	Initial	0	0	1	0
	Medial	0	0	0	1
	Final	66	30	108	88

From this analysis, it is reasonable to conclude that, by itself, tense of MC also seems not to be decisive to determine temporal relation. In effect, the same inference can be made for BE, as shown by Table 7.

Table 7. Results of the combination tense of the MC/temporal relation/position of SC in BE

Temporal relation	Position of SC	Pres	PresP	Pst
Ant	Initial	21	5	38
	Medial	7	0	8
	Final	47	9	36
Simul	Initial	0	0	0
	Medial	0	0	0
	Final	0	5	3
Post	Initial	0	0	0
	Medial	0	0	0
	Final	0	0	1

However, these results point to the conclusion that there must be a relevant difference between the perfect participle forms in PT and BE, which is pertinent to address the second problem tackled in this paper, namely the factors underlying the temporal interpretation of APC.

So, in the case of BE, when the MC has the verb in the Past, the anteriority reading of APC is almost always mandatory, irrespective of the position (cf. Table 7). This leads to the conclusion that the perfect participle in BE has a clear orientation to the past, and that information is the main source to compute the temporal location of APC in English.

Concerning PT data, when the MC has the verb in the PP, the distribution of readings is much more balanced than in BE (cf. Table 6). Furthermore, in final position, although all readings are possible, there are more posteriority readings than anteriority or simultaneity readings (except for BP). Finally, the posteriority readings in PT arise almost exclusively in the final position (only 2 exceptions in 294 cases in all varieties). All these aspects are in line with Leal's (2002, 2011) proposal that the perfect participle in PT has a temporal value of 'past' (that is, the location time is anterior to its TP_{pt}) and that this temporal feature interacts with other factors in the temporal location of APC: in particular, posteriority readings arise by combining temporal location of the main clause (typically, by using PP in the MC) with temporal information of the perfect participle ('past') and clause ordering (MC+SC).

3.3 Temporal readings and aspectual classes

Factoring in temporal relations, apparently there is no link between the MC aspectual type and the temporal relation in most of PT varieties. Though some regularities can be observed (with CP and Pro, a temporal relation of simultaneity; with Culm, of posteriority; and with St of anteriority), almost all combinatory

possibilities between aspectual types and temporal relations are perceived in the corpus. In BE, the temporal relation of anteriority prevails regardless of the MC aspectual type (see Table 8). At this point, only for AP-MP pair was no significant difference observed with a χ^2 test.

Table 8. Results of the pair aspectual type of MC/temporal relation

Aspectual type of MC/Temporal relation	EP	BP	AP	MP	BE
CP-Ant	4	6	2	3	11
CP-Simul	16	5	11	20	1
CP-Post	9	11	10	8	0
Culm-Ant	23	21	12	9	56
Culm-Simul	6	8	3	7	0
Culm-Post	54	22	93	76	0
Pro-Ant	1	5	2	0	111
Pro-Simul	21	17	18	26	6
Pro-Post	5	2	6	7	1
St-Ant	32	58	21	11	186
St-Simul	20	31	15	27	8
St-Post	8	3	5	4	3

Similar conclusions can be reached concerning the connection between the SC aspectual type and temporal relation. The differences lie in the most frequent combinatory possibilities in PT varieties: with CP, the preferential temporal interpretation varies according to the PT variety; with Culm, the tendency is a posteriority reading; with Pro and St an interpretation of simultaneity is obtained (see Table 9). Here, only for EP-AP pair no significant difference was observed through the χ^2 test.

Table 9. Results of the pair aspectual type of SC/temporal relation

Aspectual type of SC/Temporal relation	EP	BP	AP	MP	BE
CP-Ant	10	26	7	11	36
CP-Simul	14	17	14	22	0
CP-Post	12	10	15	13	1
Culm-Ant	37	43	18	9	85
Culm-Simul	30	19	20	30	0
Culm-Post	54	17	85	67	2
Pro-Ant	7	10	8	1	36
Pro-Simul	12	13	9	13	5
Pro-Post	5	4	7	7	0
St-Ant	6	11	4	2	29
St-Simul	7	12	4	15	3
St-Post	5	7	7	10	0

At first glance, this analysis seems to suggest that there is no direct link between the aspectual type of SC and MC and the temporal relation. However, when we confine the inquiry to data with the most frequent tense – PP¹³ for PT and Pst for BE – the analysis of the combination of the aspectual classes in both clauses unveils some tendencies.

When there are posteriority and anteriority relations, what seems to be more relevant in the combination of the aspectual types is the feature telicity in both clauses, the most frequent types being Culm and CP. However, with simultaneity relations, since Pro, CP and St are recurrent, it looks as if in both clauses the feature duration is more pertinent (see Table 10).

Table 10. Results of the combination aspectual type of SC/aspectual type of MC/temporal relation -EP

Temporal relation	Position of SC	MC & tense(MC) = PP				SC & tense(MC) = PP			
		CP	Culm	Pro	St	CP	Culm	Pro	St
Ant	Initial	0	2	0	0	0	2	0	0
	Medial	0	0	0	0	0	0	0	0
	Final	2	15	0	1	5	10	1	2
Simul	Initial	0	0	0	0	0	0	0	0
	Medial	0	0	0	0	0	0	0	0
	Final	16	6	13	8	10	20	6	7
Post	Initial	0	0	0	0	0	0	0	0
	Medial	0	0	0	0	0	0	0	0
	Final	8	51	4	3	11	47	4	4
	Σ	26	74	17	12	26	79	11	13
					129				129

Examples (18)–(20) give a picture of these regularities.

- (18) *Lima e o seu sócio Vitor Raposo criara em 2007 um fundo para Lima and the his partner Vitor Raposo created in 2007 a fund for esse efeito tendo pedido ao BPN um crédito de mais de this purpose, having-GER asked-PTCP to-the BPN a credit of more of 40 milhões de euros.*

40 million euros

“Lima and his partner Vitor Raposo created in 2007 a fund for this purpose, having asked BPN for a credit of over 40 million euros.”

(EP, anteriority, MC-CP; SC-Culm)

13. Recall that this is a tense that does not perform any aspectual change in the predication.

- (19) *A lei das plataformas deu entrada no parlamento em janeiro de 2017, tendo sido aprovada em março.*
 The law-of-the platforms gave entrance in-the parliament in January of 2017, having-GER been-PTCP passed in March
 “The platform law entered parliament in January 2017, having been passed in March.”
 (EP, posteriority, MC-Culm; SC-Culm)
- (20) *No entanto, foi no cinema que se destacou, dividindo-se entre o documentário e a ficção, tendo assinado mais de 50 obras.*
 However, was in-the cinema that himself stood out, dividing-himself between the documentary and the fiction, having-GER signed-PTCP more than 50 works
 “However, it was in the cinema that he stood out, dividing himself between documentary and fiction, having signed more than 50 works.”
 (EP, simultaneity, MC-St; SC-CP)

Therefore, this analysis shows that telicity triggers temporal succession, whereas durativity triggers temporal simultaneity in EP.

This tendency is also detected in the other PT varieties, as shown in Tables 11–13.

Table 11. Results of the combination aspectual type of SC/aspectual type of MC/temporal relation –AP

Temporal relation	Position of SC	MC & tense(MC) = PP				SC & tense(MC) = PP			
		CP	Culm	Pro	St	CP	Culm	Pro	St
Ant	Initial	0	1	0	1	0	1	0	1
	Medial	0	2	0	0	0	2	0	0
	Final	2	6	2	1	2	6	2	1
Simul	Initial	0	0	0	1	0	0	0	1
	Medial	0	0	0	0	0	0	0	0
	Final	10	2	11	6	10	2	11	6
Post	Initial	0	1	0	0	0	1	0	0
	Medial	0	0	0	0	0	0	0	0
	Final	10	88	6	4	10	88	6	4
	Σ	22	100	19	13	22	100	19	13
					154				154

Table 12. Results of the combination aspectual type of SC/aspectual type of MC/temporal relation –MP

Temporal relation	Position of SC	MC & tense(MC) = PP				SC & tense(MC) = PP			
		CP	Culm	Pro	St	CP	Culm	Pro	St
Ant	Initial	0	0	0	1	0	0	0	1
	Medial	1	2	0	1	1	2	0	1
	Final	2	7	0	0	2	7	0	0
Simul	Initial	0	0	0	0	0	0	0	0
	Medial	0	0	0	0	0	0	0	0
	Final	20	7	19	13	20	7	19	13
Post	Initial	0	0	0	0	0	0	0	0
	Medial	0	1	0	0	0	1	0	0
	Final	6	70	6	4	6	70	6	4
	Σ	29	87	25	19	29	87	25	19
					160				160

Table 13. Results of the combination aspectual type of SC/aspectual type of MC/temporal relation –BP

Temporal relation	Position of SC	MC & tense(MC) = PP				SC & tense(MC) = PP			
		CP	Culm	Pro	St	CP	Culm	Pro	St
Ant	Initial	1	3	4	3	3	7	1	0
	Medial	0	4	0	2	1	3	1	1
	Final	4	11	0	3	7	8	2	1
Simul	Initial	1	0	0	0	1	0	0	0
	Medial	0	0	0	1	0	0	0	1
	Final	4	7	11	18	10	13	10	7
Post	Initial	0	0	0	0	0	0	0	0
	Medial	0	0	0	0	0	0	0	0
	Final	8	19	2	1	8	12	3	7
	Σ	18	44	17	28	30	43	17	17
					107				107

Once more, BE exhibits dissimilar features (see Table 14): although the simultaneity reading involves (like in EP) durative situations – Pro and St,¹⁴ – this reading is residual, as most examples correspond to anteriority reading. Moreover, despite the fact that there are more telic situations than atelic both in MC (51 and 31,

14. There are 36 durative and 46 non-durative situations in MC and in the SC 41 durative and 40 non-durative.

Table 14. Results of the combination aspectual type of SC/aspectual type of MC/temporal relation – BE

Temporal relation	Position of SC	MC & tense(MC) = PP				SC & tense(MC) = PP			
		CP	Culm	Pro	St	CP	Culm	Pro	St
Ant	Initial	3	25	1	9	5	19	6	8
	Medial	0	3	1	4	1	5	2	0
	Final	2	18	2	14	6	16	8	6
Simul	Initial	0	0	0	0	0	0	0	0
	Medial	0	0	0	0	0	0	0	0
	Final	0	0	1	2	0	0	2	1
Post	Initial	0	0	0	0	0	0	0	0
	Medial	0	0	0	0	0	0	0	0
	Final	0	0	1	0	0	1	0	0
	Σ	5	46	6	29	12	41	18	15
					86				86

respectively) and SC (52 and 30, respectively), all aspectual types occur both in MC and SC with anteriority interpretation.

Example (21), for instance, despite the presence of two states, has a temporal reading of anteriority.

- (21) Labour hoped to retain majority control of the city chambers, *having been in power there since 1980* (BE, anteriority, MC-St; SC-St)

Several remarks are in order here concerning the relevant factors underlying temporal interpretation. The statistical analysis of the data shows that, like the ordering of clauses, the aspectual profile of the clauses has much bearing in the final interpretation of the PT APC. Conversely, in BE, the anteriority relation, seemingly, relies on the anterior orientation/perfect aspect of the perfect participle to the detriment of aspectual classes and order of the clauses.

4. Concluding remarks

The corpus-based study presented in this paper was designed to come to grips with two related problems: (i) to the distinction of two types of APC in PT based on their position and temporal interpretation, and its productivity in BE; (ii) and to the role of the temporo-aspectual nature of the perfect participle, position, tense of MC, and aspectual classes in the final temporal reading of these clauses in four PT varieties, as well as in BE.

First, it should be noted that the data provide counterevidence to Lobo's argument (2003, 2006, 2013) that peripheral clauses in PT typically occupy an initial position and have an anteriority reading, while posteriority APC, by default, are positioned at the end of the sentence and are interpreted as posterior to the main clause. Our analysis indicates that such a distinction does not explain the data, thus, failing to ensure a complete description of the APC in PT (and in BE). In point of fact, APC can clearly occupy three positions, though the most frequent are initial and final. Nonetheless, our analysis shows that those positions are not associated with a specific temporal interpretation.¹⁵

Second, the analysis of the combination of parameters points to some compelling results about the key factors behind APC temporal interpretation. According to the data, the ordering of clauses is relevant to determine the temporal relation in PT, but not in BE. Tense of MC alone apparently is not pertinent to temporal reading both in PT and BE. Aspectual classes of predications are, according to data, closely intertwined to temporal interpretation in PT (but not in BE), as evidenced by the fact that anteriority and posteriority readings tend to be related to telic situations in MC and SC, whereas simultaneity readings lean on the presence of durative situations in both clauses. However, there are no absolute restrictions regarding telicity and durativity both in MC and SC.

On the other hand, the analysis corroborates what has been reported in the literature for English in relation to the anterior orientation of APC (cf. Kortmann 1995: 221). In fact, the statistical analysis reveals that neither the position nor aspectual classes overrule the temporo-aspectual information carried by the perfect participle. This is not the case in PT. Since temporal interpretation of APC hinges on position and aspectual classes, one cannot claim that perfect participle is as determinant as it is in BE. Nonetheless, it should be said, that in addition to the anteriority readings, simultaneity and posteriority readings are also possible in BE.

One final remark regarding the comparison of PT varieties. The data show that EP, AP and MP are similar in what concerns the preferred (i) position of the APC (final); (ii) temporal meaning (posteriority); and (iii) aspectual type of MC (Culm and St), and aspectual type of the SC (Culm and CP). The similarities between EP and African varieties bear out Gonçalves's claim (2013: 161) that "the discourse of MP and AP speakers is generated not only by rules of the 'new' grammars, but also by the EP grammar". As for BP, it has several points of similarity with BE: the preferred temporal interpretation (anteriority) and the preferred aspectual type of

15. Lobo (2003) also uses as criteria to distinguish peripheral clauses from posteriority clauses the fact that the former can exhibit non-temporal values such as reason and concession and the latter only convey the temporal value. However, Silvano, Leal & Cordeiro (2019) also argue that both clauses can display a variety of rhetorical relations not predicted by Lobo.

the MC (St and Culm). In fact, the analysis shows that BP is not as similar to EP as it is suggested in the literature (cf. Neto & Foltran 2001; Mória & Viotti 2004b; Leão 2018), namely with respect to the preferred temporal meaning.

In future research, we aim to extend the analysis to more data of the corpus, and to carry out a more detailed survey of correlations between temporal relations, tenses (namely those that cause aspectual changes like *Pretérito Imperfeito* “Past with imperfective reading”), aspectual classes, position, connectors, and rhetorical relations. These next steps will provide us with a more thorough account of the semantic features of APC from a contrastive perspective at micro and macro variation levels, which may eventually lead to a typological proposal.

Acknowledgements

The authors wish to express their gratitude to the two reviewers and to Fátima Oliveira for their many comments, and to gratefully acknowledge all the work carried out by the editors. The usual disclaimers apply.

Funding

This research was supported by the Program FCT – UIDB/00022/2020 (Fundação para a Ciência e a Tecnologia).

References

- Carvalho, Ana Maria & Dante Lucchesi. 2016. Portuguese in contact. In Leo Wetzels, João Costa & Sergio Menuzzi (eds.), *The handbook of Portuguese linguistics*, 41–55. Oxford: John Wiley. <https://doi.org/10.1002/9781118791844.ch3>
- Cunha, Celso & Lindley Cintra. 1984. *Nova gramática do português contemporâneo*. Lisboa: João Sá da Costa.
- Cunha, Luís Filipe. 2004. Semântica das predicções estativas. Para uma caracterização aspectual dos estados. Ph.D. Dissertation: University of Porto.
- Cunha, Luís Filipe, António Leal & Purificação Silvano. 2015. Relações retóricas e temporais em orações gerundivas adverbiais. In Purificação Silvano & António Leal (eds.), *Estudos de semântica*, 63–71. Porto: CLUP/FLUP.
- Dowty, David. 1979. *Word meaning and Montague grammar. The semantics of verbs and times in generative semantics and in Montague’s PTQ*. Dordrecht: Reidel. <https://doi.org/10.1007/978-94-009-9473-7>
- Gonçalves, Perpétua. 2010. *A génese do português de Moçambique*. Lisboa: INCM.
- Gonçalves, Perpétua. 2013. O português em África. In Eduardo Paiva Raposo, Maria Fernanda Bacelar do Nascimento, Maria Antónia Mota, Luísa Segura & Amália Mendes (eds.), *Gramática do português*, Volume 1, 157–178. Lisbon: Fundação Calouste Gulbenkian.

- Haspelmath, Martin. 1995. The converb as a cross-linguistically valid category. In Martin Haspelmath & Ekkehard König (eds.), *Converbs in cross-linguistic perspective: Structure and meaning of adverbial verb forms – adverbial participles, gerunds* (EALT, 13), 1–55. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110884463-003>
- Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic. Introduction to modeltheoretic semantics of natural language, formal logic and discourse representation theory*. Dordrecht: Kluwer.
- Klein, Dan & Christopher D. Manning. 2003. Accurate unlexicalized parsing. In Erhard W. Hinrichs & Dan Roth (eds.), *Proceedings of the 41st annual meeting on Association for Computational Linguistics*, 1, 423–430. Stroudsburg, PA: Association for Computational Linguistics. <https://doi.org/10.3115/1075096.1075150>
- König, Ekkehard. 1995. The meaning of converb constructions. In Martin Haspelmath & Ekkehard König (eds.), *Converbs in cross-linguistic perspective. Structure and meaning of adverbial verb forms – adverbial participles, gerunds*, 1–56. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110884463-004>
- Kortmann, Bernard. 1995. Adverbial participial clauses in English. In Martin Haspelmath & Ekkehard König (eds.), *Converbs in cross-linguistic perspective. Structure and meaning of adverbial verb forms – adverbial participles, gerunds*, 189–238. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110884463-007>
- Leal, António. 2002. O valor temporal das orações gerundivas em português. In Amália Mendes & Tiago Freitas (eds.), *Actas do XVIII Encontro da Associação Portuguesa de Linguística*, 455–464. Porto: APL.
- Leal, António. 2009. Semântica aspectual e nominal. Contributo das expressões nominais para a construção aspectual das frases. Ph.D. Dissertation. Porto: University of Porto.
- Leal, António. 2011. Some semantic aspects of gerundive clauses in European Portuguese. In Jesse Mortelmans, Tanja Mortelmans & Walter De Mulder (eds.), *From now to eternity. Cahiers Chronos*, Volume 22. 85–113. Amsterdam: Editions Rodopi.
- Leão, Rafaella. 2018. A semântica das construções gerundivas no português europeu e no português do Brasil. Ph.D. Dissertation: Universidade do Porto.
- Lobo, Maria. 2003. Aspectos da sintaxe das orações subordinadas adverbiais. Ph.D. Dissertation: Universidade Nova de Lisboa.
- Lobo, Maria. 2006. Dependências temporais: A sintaxe das orações subordinadas gerundivas do Português. *Veredas* 10(1–2).
- Lobo, Maria. 2013. Subordinação adverbial. In Eduardo Paiva Raposo, Maria Fernanda Bacelar do Nascimento, Maria Antónia Mota, Luísa Segura & Amália Mendes (eds.), *Gramática do português*, Volume 2, 1986–2030. Lisboa: Fundação Calouste Gulbenkian.
- López, Vanessa. 2019. The non-anteriority compound gerund and its licensing conditions. MA Thesis: Universidade de Lisboa.
- Moens, Marc. 1987. Tense, aspect and temporal reference. Ph.D. Dissertation: University of Edinburgh.
- Móia, Telmo & Evani Viotti. 2004a. Sobre a semântica das orações gerundivas adverbiais. In Inês Duarte & Isabel Leiria (eds.), *Actas do XX encontro da Associação Portuguesa de Linguística*, 715–729. Lisboa: Associação Portuguesa de Linguística.
- Móia, Telmo & Evani Viotti. 2004b. Differences and similarities between European and Brazilian Portuguese in the use of the «gerúndio». *Journal of Portuguese Linguistics* 3(1). 111–139. <https://doi.org/10.5334/jpl.21>

- Muniz, Marcelo C., Maria das Graças V. Nunes & Eric Laporte. 2005. UNITEX-PB, a set of flexible language resources for Brazilian Portuguese. In Rove Chishman (ed.), *Proceedings of Workshop on technology on information and human language (TIL)*, 2059–2068. São Leopoldo, Brazil: Universidade do Vale do Rio dos Sinos.
- Najork, Marc. 2009. Web crawler architecture. In Ling Liu & M. Tamer Özsu (eds.), *Encyclopedia of database systems*, 3462–3465. New York: Springer.
https://doi.org/10.1007/978-0-387-39940-9_457
- Neto, José Borges & Maria José Foltran. 2001. Construções com gerúndio. In Clara Nunes Correia & Anabela Gonçalves (eds.), *Actas do XVI encontro Nacional da Associação Portuguesa de Linguística*, 725–735. Lisbon: Associação Portuguesa de Linguística.
- Oliveira, Fátima. 2013. Tempo verbal. In Eduardo Paiva Raposo, Maria Fernanda Nascimento, Maria Antónia Mota, Luísa Segura & Amália Mendes (eds.), *Gramática do Português*, 509–556. Lisbon: Fundação Calouste Gulbenkian.
- Olston, Christopher & Marc Najork. 2010. Web crawling. *Foundations and Trends® in Information Retrieval* 4(3). 175–246. <https://doi.org/10.1561/1500000017>
- Peres, João. 1994. Sobre a Semântica das Construções Perfectivas do Português. In Inês Duarte & Isabel Leiria (eds.), *Actas do Congresso Internacional sobre o Português* (Volume 2), 33–58. Lisbon: APL & Edições Colibri.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech & Jan Svartvik. 1985. *A comprehensive grammar of the English language*. London: Longman.
- Risvik, Knut Magne & Rolf Michelsen. 2002. Search engines and web dynamics. *Computer Networks* 39(3). 289–302. [https://doi.org/10.1016/S1389-1286\(02\)00213-X](https://doi.org/10.1016/S1389-1286(02)00213-X)
- Sæbø, Kjell Johan. 2011. Adverbial clauses. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning*, Volume 2, 1420–1441. Berlin: De Gruyter Mouton.
- De Smet, Hendrik. 2010. English *ing*-clauses and their problems: the structure of grammatical categories. *Linguistics* 48(6). 1153–1193. <https://doi.org/10.1515/ling.2010.038>
- Silvano, Purificação, António Leal & João Cordeiro. 2019. Algumas reflexões sobre a classificação de orações gerúndivas em português europeu. *Revista da Associação Portuguesa de Linguística* 5. 325–347. <https://doi.org/10.26334/2183-9077/rapln5ano2019a22>
- Stump, Gregory T. 1985. *The semantic variability of absolute constructions*. Dordrecht: Reidel.
<https://doi.org/10.1007/978-94-009-5277-5>

Craindre (“fear”) and expletive negation in diachrony

Chloé Tahar

Institut Jean Nicod, DEC, ENS, EHESS, CNRS, PSL University

This paper investigates the distribution of expletive negation in the complement clause of *craindre* (“fear”) in French. Building on Anand & Hacquard’s (2013) proposal that fear verbs are hybrid attitude verbs, featuring both a doxastic and a (dis)preferential component, this paper argues that these two components are conveyed by different layers of meaning (in line with Giannakidou & Mari (2020)). More precisely, I argue that, in actual discourse context, *craindre* may receive two main interpretations: a volitive (dispreference-related) or a psychological (belief-related) interpretation, depending on whether the verb *asserts* or *presupposes* dispreference. Based on a diachronic corpus study of the distribution of expletive negation, I show that expletive negation, in the earliest stages of French, places semantic restrictions on the main verb, which are met when the interpretation of *craindre* is volitive.

Keywords: expletive negation, modality, mood, diachronic semantics, French

1. Introduction

Expletive negation is the form-meaning mismatch whereby a negation marker (*ne* in French) is not interpreted as such in certain contexts, receiving an ‘expletive’, or semantically colorless, reading. Previous approaches to the issue of expletive negation have revolved around the idea that occurrences of expletive negation are restricted to so-called *adversative* contexts: contexts which share an abstract negative semantic property. One of the most sophisticated proposals within this perspective are those of Muller (1991) and van der Wurff (1999). These authors argue that the set of adversative predicates that trigger expletive negation are negative insofar as they can be paraphrased by a corresponding ‘positive’ verb embedding a negated proposition (e.g., fear that p = wish that not p). In the spirit of Jespersen’s

(1917) seminal proposal that expletive negation is in fact a ‘paratactic’ negation in Latin, introducing its own independent clause, it has later been argued that the semantic colorlessness of expletive negation results from its redundancy with the main verb’s ‘silent’ lexical negation. For instance, Parry (2013) defends the idea that paratactic negation could have historically lost its inherent semantic negativity as a consequence of tighter syntactic subordination to the main verb, and thus increased redundancy. Empirical studies have supported this hypothesis, showing that expletive negation in diverse Indo-European languages such as Greek, Albanian or Latin originates from the negation marker dedicated to the construction of prohibitive clauses, see Mari & Tahar (2020). Such studies allow for the stipulation that prohibitive negation in those languages underwent the following developmental pathway: May it not be true! > I fear: may it not be true! > I fear that it may (not) be true. The present paper builds on the claim that *ne* was a Latin speech-act negator that developed into an expletive negation (in line with Ageno (1955); Lakoff (1968); Chatzopoulou (2012); Parry (2013), and Mari & Tahar (2020)).

The goal of this paper is to address one of the important limitations of the adversative hypothesis. As the adversative hypothesis takes the expletive negation marker to somehow spell out the main verb’s lexical negation, it cannot account for the syntactic and semantic factors underpinning the presence (rather than absence) of expletive negation in the predicate’s subordinate clause. The present paper aims to address this issue by identifying differences in the meaning of *craindre* with and without expletive negation in its complement clause. Our starting point is a distributional puzzle often noticed in the French literature (see Gaatone (1971); Larrivée (1994)): the anti-licensing of expletive negation in the complement clause of the negated form of *craindre*.

- (1) a. *Je crains que vous ne soyez en retard encore une fois.*
 I fear.1SG.IND that you EN be.2SG.SBJV in late again one time
 “I’m afraid you are late once again.”
- b. *Je ne crains pas que vous (#ne) soyez en retard encore une fois.*
 I NEG fear.1SG.IND not that you EN be.2SG.SBJV in late again
 one time
 “I’m not afraid you are late once again.”

As observed by Larrivée (1994), when *craindre* is under negation as in (1–b), it conveys the subject’s negative belief with respect to the possibility that *p*. In other words, in Example (1–b), the speaker expresses disbelief with respect to the possibility that the addressee will ever be late again. Importantly, a subtle ambiguity can show up in the meaning of *craindre* under negation: it may as well express the subject’s dispreference denial; see the contrast between (2) and (3).

- (2) (A hypochondriac patient goes to see the doctor because he believes that he suffers from the polio virus. The doctor finds out that the patient is perfectly healthy.)
- a. *Je ne crains pas que vous (#n') ayez la polio.*
 I NEG fear.1SG.IND not that you EN have.2SG.SBJV the polio
 “I don’t fear that you have the polio virus (you are in good shape).”
- (3) (A banker, suspected of fraud, is ordered to show his account books, which he does.)
- a. *Je ne crains pas que vous ne voyiez mes livres de comptes.*
 I NEG fear.1SG.IND not that you EN see.2PL.SBJV my books
 of accounts
 “I don’t fear that you see my books of accounts (I’ve got nothing to hide).”

In (2), *craindre* under negation has the same meaning it has in (1–b) – that of expressing disbelief – and anti-licenses expletive negation. In (3), however, *craindre* under negation expresses the banker’s denial that he has a dispreference for his books of accounts to be openly checked. The possibility that his books of accounts will be openly checked is nonetheless very likely, if not factual.

It’s only when *craindre* under negation expresses dispreference denial that it can license expletive negation. The contrast between (2) and (3) could be taken to indicate that expletive negation requires the main predicate to express the subject’s positive belief with respect to the possibility that *p* in order to be licensed, as argued by authors like Espinal (2007) and Yoon (2011). This paper argues that the contrasted behavior of expletive negation in the complement clause of *craindre* under negation is also the manifestation of an ambiguity rooted in the verb’s semantics.

Adopting Anand & Hacquard’s (2013) proposal that fear verbs are hybrid attitude predicates, featuring both a doxastic and a (dis)preferential component, the core meaning of *craindre* is as given in (4).

- (4) *a* craint que *p*:
- a. **Doxastic layer:** there is a *w*’ in *Dox(a, w)* such that *p* is true in *w*’
 b. **Dispreference layer:** for every *w*’ in *Bul(a, w)*, not-*p* is true in *w*’

The doxastic component of *craindre* is captured as existential quantification over *p* belief worlds. This doxastic layer expresses belief with respect to the truth of the proposition *p*, as well as uncertainty (see Giannakidou (1999); Giannakidou & Mari (2017)). For simplicity, the semantics given above deals with dispreference as universal quantification over not-*p* desire worlds.

In the line of reasoning of authors like Farkas (1985); Giannakidou & Mari (2016), and Giannakidou & Mari (2020), I will argue that these modal components are conveyed at two distinct levels of interpretation. More specifically, I will argue

that *craindre* receives two main types of interpretations depending on whether it conveys dispreference at the level of the *assertion* or rather at the level of the *presupposition* (Paul Portner p.c). The first type of interpretation, labelled ‘psychological’, foregrounds the doxastic component of the verb, as in (5). The preferential layer is thus backgrounded.

- (5) (Calling out the neighbour for information)

Je crains qu’ il soit arrivé quelque chose à mon chien.
 I fear.1SG-IND that it be.3SG.SBJV happened some thing to my dog
 “I fear that something happened to my dog (I haven’t seen it this morning).”

Psychological *craindre* asserts the speaker’s uncertainty. It raises the question of the truth or falsity of the prejacent *p* and thus indirectly constitutes a request for information. When psychological *craindre* is under negation, as in (2), repeated below as (6), it asserts disbelief, presupposes dispreference, and does not felicitously trigger expletive negation.

- (6) *Je ne crains pas que vous (#n’) ayez la polio.*

I NEG fear.1SG.IND not that you EN have.2SG.SBJV the polio
 “I don’t fear that you have the polio virus (you are in good shape).”

- (7) *a ne craint pas que p:*

- a. **Disbelief assertion:** there is no *w*’ in *Dox(a, w)* such that *p* is true in *w*’
- b. **Dispreference presupposition:** for every *w*’ in *Bul(a, w)*, not-*p* is true in *w*’

Contrary to the psychological interpretation of *craindre*, the second type of interpretation of *craindre*, labelled ‘volitive’, foregrounds the dispreference component of the verb. It asserts the speaker’s dispreference with respect to a possible situation in the future course of events, see (8), while the doxastic layer is backgrounded.

- (8) (Up on a hill, to a swimmer entering a zone full of jellyfishes)

A: *Attention! Je crains que vous ne vous fassiez piquer par*
 Watch-out! I fear.1SG.IND that you EN you do.2PL.SBJV sting by
une méduse!
 a jellyfish
 “Watch out! I fear you will get stung by a jellyfish!”

The assertion is not meant to be challenged by the addressee on epistemic grounds, (for instance, with a request for the justification of the belief) but rather to be complied with or rejected. The addressee of (8) may (or not) comply with the indirect command to change his trajectory and go back to the shore in a hurry. When volitive *craindre* is under negation, it denies dispreference (i.e., is a case of ‘illocutionary denegation’ à la Searle (1969)), presupposes belief, as in (3), repeated below, and may felicitously trigger expletive negation.

- (9) *Je ne crains pas que vous ne voyiez mes livres*
 I NEG fear.1SG.IND not that you EN see.2PL.SBJV my books
de comptes.
 of accounts
 “I don’t fear that you see my account books (I’ve got nothing to hide).”

- (10) *a ne craint pas que ... ne p:*
 a. **Dispreference denial:** \sim ASSERT(every w' in $\text{Bul}(a, w)$ is such that not- p is true in w')
 b. **Doxastic presupposition:** there is a w' in $\text{Dox}(a, w)$ such that p is true in w'

In this paper, I explore the hypothesis that expletive negation is first and foremost sensitive to the volitive reading of *craindre*. To this end, a diachronic corpus study has been conducted, which takes a closer look at the diverse interpretations of the verb in actual discourse context. This corpus study aims to examine the influence of expletive negation on the verb’s interpretation. Section 2 describes three possible uses of *craindre* and draws hypothetical correlations between the presence or absence of expletive negation in the complement and the pragmatic purpose of the assertion of *craindre*. Section 3 presents the results of the study, where these correlations are examined in a diachronic perspective. Section 4 presents an analysis for the distribution of expletive negation with the three uses of *craindre*. Section 5 concludes.

2. Psychological and volitive readings of *craindre*

This section provides a description of the different uses that *craindre* may receive in actual discourse context (as advocated by Yanovich (2013) among others), based on a fine-grained semantic analysis of examples from the Frantext corpus (to be introduced in the next section). In this section, I describe the prototypical uses that *craindre* may receive, in the presence or absence of expletive negation in its complement. I argue that *craindre* mostly receives a psychological reading with the mere complementation form *que*, as evidenced by the fact that the assertion of *craindre que* in context serves an inquisitive (belief-related) purpose. When *craindre* takes the complementation form *que ... ne*, it mostly receives a volitive reading, as evidenced by the fact that its assertion may either serve an admonitive or reprehensive purpose (dispreference-related purposes). Note that the three uses singled out (inquisitive, admonitive and reprehensive) do not exhaust all possible uses of *craindre*. These are rather prototypical correlations observed between the range of uses that *craindre* may receive and its complementation forms.

2.1 Inquisitive *craindre que*

Practical goal

The assertion of *craindre que* has an inquisitive goal when used to make a conjecture, as in Example (11), (12). The speaker conjectures that *p* if s/he commits herself to the belief that *p* in a way which is liable to error.

- (11) *Moy (sotte) qu'ay-je à craindre ? ... D'où me vient ceste crainte? Je crain qu'amour m'en soit la cause.*

“Me (fool) what do I have to fear? Where does this fear come from? I fear that love may cause me this fear.” (Louis Des Mesures, David Triomphant, 1566)

- (12) (Eugène confesses that he suspects Alix of cheating on him).
Mon amour est douteuse : et je crains que cette mignarde d'aller autre part se hasarde.

“My love is suspicious as I fear that this cute thing is seeing someone else.”
(Etienne Jodelle, L'Eugène, 1573)

Presupposition projection under negation

Inquisitive *craindre que* under negation foregrounds disbelief, expressing that the attitude holder doesn't take the proposition *p* to be possible. In (13), the sentence expresses that the attitude anchor doesn't believe that the sky will fall. In (14), the sentence expresses that the attitude anchor doesn't believe that her parents will scold her. In both sentences, the attitude anchor's dispreference for *p* is conventionally implied.

- (13) *Oh! Je ne crains pas que le ciel me tombe sur la tête!*

“Oh! I don't fear that somehow the sky would fall!”

(Jules Vernes, Voyage au centre de la Terre, 1864)

- (14) *Elle ne craignait pas que nous la grondions (car nous ne la grondions jamais).*

“She didn't fear that we would scold her (because we never did).”

(Philippe Forest, Toute la nuit, 1999)

I thus observe that superordinate negation here targets the doxastic layer of the verbs' semantics and leaves the dispreference component untouched, as represented in (15).

- (15) *a* ne craint pas que *p*:

- a. **Disbelief assertion:** there is no w' in $\text{Dox}(a, w)$ such that *p* is true in w'
- b. **Dispreference presupposition:** for every w' in $\text{Bul}(a, w)$, not-*p* is true in w'

Mood selection behavior

With the complementation form *que*, *craindre* patterns with *espérer* (“hope”), as it could select for the indicative mood from 1100 to 1650 (27 occurrences in total occurred in the corpus). Note that this is not uncommon on a cross-linguistic perspective, since, as observed by Farkas (1992), *fear* and *hope* obligatorily select for the indicative in Romanian.

- (16) *Je crains que c'est.IND un traître.*
 “I fear that he’s a traitor.” (Michel de Montaigne, *Les Essais*, 1582)
- (17) *Si tu ne sais pas mieux donner remède à mon mal, j’ay grand peur que j’en seray. IND longuement malade.*
 “If you cannot cure my disease, I fear that I will remain sick for a long time.”
 (Honoré d’Urfé, *L’Astrée*, 1612)

The ability of *craindre que* to select for the indicative at a certain period of the history of French is revelatory of the fact that its assertion conveys the belief component of the verbs’ semantics. It has indeed been argued that the belief layer of attitude verbs is responsible for indicative mood choice (see also Farkas (1985); Giorgi & Pianesi (1997); Giannakidou (1998); Villalta (2008); Portner & Rubinstein (2012); Anand & Hacquard (2013)).

2.2 Admonitive *craindre que ... ne*

Practical goal

The assertion of *craindre que ... ne* has an admonitive goal when used to express a warning, see (18), (19). When a sentence with *craindre* in it has a flavor of warning, the speaker expresses a public dispreference for the truth of *p* in the future course of events (implying the possibility of taking some precautions to avoid this possible future situation, see Lichtenberk (1995)). The addressee may thus draw the inference that s/he’s requested to bring about the falsity of *p*. Note that, in Old French, admonitive uses of *craindre* often arise in the context of supplication for the intervention of an absent (and divine) addressee, to keep the speaker safe from danger, as in (18).

- (18) (The Lord begs the Virgin to help him against the Devil’s temptations)
LE SEIGNEUR. Se vous ne m’aidiez sanz attente, je crain que je n’y soie atains; car je suis seur et certains, Vierge, que il me suit et gaité.
 “THE LORD. If you don’t help me immediately, I fear that I’ll get caught; because I’m absolutely certain, Virgin, that he’s watching out for me.”
 (Anonymous, *Miracle de l’enfant donné au diable*, 1339)

(19) (The speaker is dying)

Hâte-toi, mon Georges, hâte-toi, je crains que tu n'arrives trop tard.

“Hurry on, my Georges, hurry on, I fear that you will be late.”

(Léon Bloy, *Le désespéré*, 1886)

Presupposition projection

In the Frantext corpus, no occurrence was found of expletive *ne* in the complement of *craindre* under negation, except for one occurrence, given in (20).

(20) (Phillis denies that she doesn't want Alexis to join her, Diane and Astrée, but makes it clear that she doesn't want Astrée to grow too close to Alexis)

Je ne crains pas que nous ne l'emmenions [Alexis] Mais je voy desja ... que ceste Astrée nous quittera pour ceste nouvelle venue.

“I do not fear that we take Alexis with us However, I can foresee ... that our Astrée will leave us for this newly arrived Alexis.”

(Honoré d'Urfé, *L'Astrée*, 1612)

Sentence (20) conveys that Phillis denies having a dispreference for Alexis to come with her, Diane and Astrée. When *craindre* under negation selects for a complement with expletive negation, it thus denies the subject's dispreference for *p*, while presupposing the subject's positive belief with respect to the likelihood of *p*, as represented in (21).

(21) *a ne craint pas que ... ne p*

a. **Dispreference denial:** \sim ASSERT(every w' in $\text{Bul}(a, w)$ is such that not- p is true in w')

b. **Doxastic presupposition:** there is a w' in $\text{Dox}(a, w)$ such that p is true in w'

Note that the scarcity of sentences where *craindre* under negation triggers expletive negation in the Frantext corpus can be explained by the fact that the most natural use of *craindre* under negation is to assert disbelief.

Mood selection behavior

In French, expletive negation is, at least in attitudinal context, always correlated to the subjunctive mood. The sensitivity of expletive negation to the subjunctive mood is illustrated below with the verb *suspecter* (“suspect”), which allows for expletive negation on the one hand and for mood shift on the other, but not for expletive negation along with the indicative mood, see (22b).

(22) a. *Ils suspectent que Jean ne soit.SBJV le meurtrier.*

“They suspect that Jean is the murderer.”

b. *Ils suspectent que Jean (#n') est.IND le meurtrier.*

“They suspect that Jean is the murderer.”

The sensitivity of expletive negation to the subjunctive mood highlights the fact that its presence in the complement of *craindre* foregrounds the verb’s dispreference layer. Indeed, preferential semantics have been argued to license the subjunctive (see also Farkas (1985); Giorgi & Pianesi (1997); Villalta (2008); Portner & Rubinstein (2012); Anand & Hacquard (2013)).

2.3 Reprehensive *craindre que ... ne*

Practical goal

Craindre que ... ne can also be used with a more specific, reprehensive goal when used to express a reproach, as in (23), (24), or an apology, and thus to express a counterfactual negative preference for the prejacent *p*.

- (23) (Martial reproaches his brother with being lazy)
Ma mère avait raison ... je crains que tu n’aies du vice ...
 “My mother was right ... I’m afraid you are vicious.”
 (Eugène Sue, *Les mystères de Paris*, 1843)

- (24) (Emile seeks to humiliate Constantin Galuchet)
Je crains que vous n’ayez fait trop boire M. Constant Galuchet. Voyez donc comme il a les yeux rouges et le regard fixe!
 “I’m afraid you pushed M. Constantin Galuchet to overdrink. Look how red and vacant-eyed he is!” (Sand, George, *Le péché de M. Antoine*, 1845)

Compatibility with the speaker’s knowledge that p

The reprehensive use of *craindre* presupposes strong belief. This use is accessible when the uncertainty presupposition encoded by *craindre* is violated and the speaker knows that the prejacent is true, unlike prototypical uses of *craindre* (see also Anand & Hacquard (2013)).

- (25) It’s raining.
 a. *Je crains qu’il ne pleuve.* (“I’m afraid it’s raining.”)
 b. *#Je crains qu’il pleuve.* (“I fear that it’s raining.”)

3. Corpus

The Frantext corpus covers periods corresponding to (Late) Old French (1100–1300), Middle French (1300–1550), Preclassical French (1550–1650), Classical French (1650–1800) and Modern French (1800–1950). The number of literary texts accessible for each period of the Frantext corpus are not homogeneous, as most ancient texts are scarcer than modern day texts. More concretely, the corpus has 28 texts from the 12th century, 30 texts from the 13th century, 115 texts from

the 14th century, 160 texts from the 15th century, 188 from the 16th century, 633 from the 17th century, 701 from the 18th century and 1150 from the 19th century. Occurrences of *craindre* with 1st person singular subject and finite complementation were retrieved within their larger context (700 words) for each period of the corpus. These sentences were first classified with respect to their complementation type (mere *que* vs. *que ... ne*). They were then manually annotated by the author as either belonging to the category of assertions of *craindre* with an inquisitive, admonitive or reprehensive goal.

Table 1. Interpretations of *craindre* with the mere complementizer *que*

Period	1st person uses	Inquisitive	Admonitive	Reprehensive	Number of words
Old Fr.	6 (35%)	5	0	1	3 388 302
Middle Fr.	2 (11%)	2	0	0	9 776 478
Preclassical Fr.	62 (34%)	44	14	4	18 994 976
Classical Fr.	14 (4%)	11	3	0	49 820 596
Modern Fr.	16 (4%)	10	3	3	76 605 469

Table 2. Interpretations of *craindre* with *que ... ne*

Period	1st person uses	Inquisitive	Admonitive	Reprehensive	Number of words
Old Fr.	11 (65%)	0	10	1	3 388 302
Middle Fr.	16 (89%)	1	14	1	9 776 478
Preclassical Fr.	122 (66%)	26	77	19	18 994 976
Classical Fr.	338 (96%)	227	63	48	49 820 596
Modern Fr.	404 (96%)	291	43	70	76 605 469

Overall, the form of complementation *que* represents a significant share of the 1st person uses of *craindre* sentences in the three first periods of the corpus (from 1100 to 1650), before growing very weak by 1650, in comparison to the form *que ... ne*. Besides, assertions of *craindre* with mere *que* are mostly associated with the inquisitive use in all the different periods of the corpus. Concurrently, for each period of the corpus, the share of assertions of 1st person *craindre* with the complementation form *que ... ne* is higher than that of *craindre* sentences with the mere complementizer *que*. Expletive negation becomes even more productive with *craindre* by Preclassical French (by 1650). Interestingly, *craindre que ... ne* doesn't receive the same use throughout the whole corpus. It is predominantly used with an admonitive goal in the three first periods of the corpus (from 1100 to 1650). By Classical French, the share of inquisitive and reprehensive uses of *craindre que ... ne* strongly increases, as the share of admonitive uses recedes.

4. Diachronic analysis

4.1 The original semantics of *ne*: A speech-act negation

In Latin, *ne* introduces an embedded negative imperative clause on the mode of quasi-parataxis. Such a structure can still be found in Early Old French, as shown in (26) and (27).

- (26) (Saint Thomas Becket advises the King of England not to conduct a policy hostile to the clergy.)

Mais cil qui tuz li munz ne remue ne tente, jo criem, Sire, ne turt. Mais ja Deus nel consente !

“But the one who is not affected nor tempted by the world, I fear, Sir, he will go away. But may God never allow that!”

(Guernes de Pont-Sainte-Maxence, *Vie de Saint Thomas Becket*, 1173)

- (27) (Robin’s master, John, is going through great suffering because of his love for Blonde. Blonde asks Robin what illness John has.)

“*Dame*”, *dist il*, “*bien le savés, pour noiant enquis le m’avés. Bien savés la mort ki le touce. Je criem Dix ne le vous reproche.*”

““Lady”, he says, “you know well; you have asked me to no purpose. You well know the death that is touching him. I fear that God will reproach you for this.””

(Anonymous, *Jehan et Blonde*, 1240)

As a speech-act negation, I argue that *ne* places semantic restriction on the utterance context. Indeed, for the utterance of a negative imperative to be felicitous, the likelihood of the proposition *p* must be available in the conversational context, as shown with the contrast between (28) and (29).

- (28) (To my friend who loves eating cookies.)

Don’t eat the last cookie!

- (29) (To my friend who is allergic to eating cookies.)

**Don’t eat the last cookie!*

The belief condition that restrains the utterance of negative imperatives is violated in (29) and, as a consequence, (29) cannot fulfill the function of a negative imperative (here, that of prohibiting). I argue that prohibitive negation, in embedded imperative clauses, places a semantic restriction on the embedding context, thus activating the possibility that *p* (see also Ducrot (1985); Dryer (1996); Larrivée (2010)) at the level of presupposition, while asserting a preference for not-*p*.

- (30) [[**ne** *p*]] is defined iff the embedding context conveys:

- Doxastic presupposition:** there is a *w*’ in *Dox*(*a*, *w*) such that *p* is true in *w*’
- If defined, [[**ne** *p*]] = for every *w*’ in *Bul*(*a*, *w*), not-*p* is true in *w*’

As quasi-paratactic clauses ultimately develop into subordinate clauses, by Late Old French, *ne* can no longer occur without the complementizer *que*, the use of which is generalized. From Late Old French to the end of Preclassical French, I argue that expletive negation retains its original semantics as a Latin speech-act negator and thus constrains the interpretation of *craindre*. Indeed, from 1100 to 1650, there is a correlation between the presence of expletive negation and the admonitive interpretation of *craindre*, the purpose of which is to warn the addressee, see (31) and (32).

- (31) *S'or ne m'en fui, molt criem que ne t'en perde.*
 “If I don’t escape now, I fear that you will go through some trouble.”
 (Anonymous, *La vie de Saint-Alexis*, 1040)
- (32) (A merchant advises Bérinus to follow his advice and accept bargaining with him)
“Sire, par ma foy, volentiers en feray par vostre conseil, mais mout crain que je ne mesface.”
 “Sir, I swear that I would like to do as you advise me to, but I strongly fear that I will do wrong.”
 (Anonymous, *Roman de Bérinus*, 1350)

With the mere complementizer *que*, the assertion of *craindre* has a different function. It describes the speaker’s epistemic state of uncertainty about whether or not the state of affairs described by *p* is the case, and the sentence is used as an indirect request for information as in (33).

- (33) (A damsel confesses that she has never loved any man except Gauvain. Having given him her virginity, she wonders if Gauvain has lost interest in her.)
Car vos estes de tel renon que je vos ai amé pieç’a. Si criem que mon domaige i a, einsin que vos ne m’amez mie.
 “For you have so great a reputation that I have loved you for a long time. Therefore I fear that I might have been treated unfairly and that you do not love me.” (Wauchier de denain, *Deuxième continuation de Perceval*, 1210)
- (34) *Je criendroie que m’en tigniez pour prinsaltiere se vous mandoie amor premiere.*
 “I would fear that you would find me impulsive if I asked for your love beforehand.”
 (Anonymous, *Le roman d’Eneas*, 1160)

In this perspective, the correlation between the presence of expletive negation and the admonitive use of *craindre* in the first three periods of the corpus is to be related to the fact that this use of *craindre* satisfies the semantic restriction that speech-act-like *ne* places on its embedding context.

- (35) *a* crait-ADMONITIVE *que ... ne p*
 a. **Dispreference assertion:** for every w' in $\text{Bul}(a, w)$, $\text{not-}p$ is true in w'
 b. **Doxastic presupposition:** there is a w' in $\text{Dox}(a, w)$ such that p is true in w'

4.2 The gradual loss of *ne*’s speech-act potential

By the last two periods of the corpus, in Classical and Modern French, assertions of *craindre* without expletive negation remain associated with an inquisitive goal.

- (36) *Mon fils bien-aimé est parti ... et je pleure sur lui. Je crains qu’il souffre chez le peuple où il est allé.*
 “My beloved son is gone ... and I’m crying over him. I fear he might suffer among the people he’s gone to.”
 (Leconte de Lisle, *Odyssée*, 1868)

Surprisingly, assertions of *craindre* with expletive negation are increasingly associated with an inquisitive goal as well, as the main point of the assertion with *craindre que ... ne* is to express the speaker’s uncertainty about whether or not the complement is true, see (37) and (38).

- (37) *Votre carte ... était bordée de deuil, et je crains que vous n’ayez eu la douleur de perdre votre sœur. Si je me trompe, comme je veux l’espérer*
 “Your letter ... was surrounded by grief, and I fear you have had the pain of losing your sister. If I’m mistaken, as I wish I am”
 (Mallarmé, *Correspondance*, 1975)
- (38) *Je crains que Mam’zelle Choute ne me soupçonne d’avoir volé cet argent. On est tellement soupçonneux à l’égard des enfants noirs.*
 “I fear that Mam’zelle Choute may suspect me of having stolen that money. People tend to be so suspicious towards black children.”
 (Joseph Zobel, *La rue Case-Nègres*, 1950)

I argue that this fact follows from the gradual path of semantic change that *craindre* underwent. Following a robust cross-linguistic tendency (see Lichtenberk (1995)), the semantics of *craindre* weakened from priority modality (i.e., the kind of modality expressing deontic, teleological or buletic (dis)preference, see also Portner (2009)) to epistemic modality. This trajectory, I argue, consists in the fact that the belief component of *craindre*, which attributes a certain psychological state of belief to the attitude anchor, moves to the foreground, while dispreference is backgrounded.

- (39) *a* craint-INQUISITIVE *que ... (ne) p*
 a. **Doxastic assertion:** there is a w' in $\text{Dox}(a, w)$ such that p is true in w'
 b. **Dispreference presupposition:** for every w' in $\text{Bul}(a, w)$, not- p is true in w'

This path of semantic change is reflected by the temporal orientation of the attitude, as displayed in Table 3. It has been argued (see for instance Condoravdi (2002)) that the temporal orientation of modals reflects the kind of modality they express, as certain types of preference-related attitudes are closely related to future-orientation.

What the data shows is that in the three first periods of the corpus, *craindre* was a future-oriented attitude, expressing a possibility about a state of affairs located in the future (thus a possibility upon which the addressee can potentially take action). It also shows that the temporal orientation of *craindre* gradually shifts towards the past, expressing a possibility about a state of affairs located in the past or in the present (which is thus metaphysically settled), by the last two periods of the corpus.

Table 3. Distributions of the temporal orientation of *craindre*

Period	1st person uses	Future	Non-future
Old Fr.	16	16	0
Middle Fr.	18	15	3
Preclassical Fr.	184	132	52
Classical Fr.	352	207	145
Modern Fr.	420	136	284

Importantly for the current discussion, as *craindre* shifts towards an epistemic meaning, the verb also develops a pragmatically specialized interpretation. By the last two periods of the corpus, indeed, *craindre* develops a reprehensive interpretation, arising with certain specific contextual factors to yield a flavor of reproach (or apology). Note that the reprehensive interpretation is almost exclusively available with the complementation form *que ... ne*.

- (40) *Je crains, docteur, que vous n'abusiez un peu de ma crédulité. Prenez-y garde.*
 "I'm afraid, doctor, you are abusing my trust. Watch out."
 (Diderot, Denis, *Le rêve de d'Alembert*, 1784)
- (41) *LE VICE-ROI : Je suis très mécontent de vous. De tous côtés on parle de votre coquetterie, et, s'il faut parler net, je crains que vous ne me fassiez jouer un sot rôle.*
 "THE VICE-ROY: I am very dissatisfied with you. Comments about your coquetry are coming from all sides. To put it bluntly, I'm afraid you are making a fool of me."
 (Mérimée, Prosper, *Théâtre de Clara Gazul*, 1857)

I propose that this specialized interpretation of *craindre* grows stronger by the last two periods of the corpus because, with this use, *ne* retains its speech-act potential. Indeed, reprehensive-*craindre* satisfies the semantic restrictions that speech-act *ne* places on the context, as it presupposes (strong) belief.

- (42) *a* craint-REPREHENSIVE *que ... ne p*
- Dispreference assertion:** for every w' in $\text{Bul}(a, w)$, not- p is true in w'
 - Doxastic presupposition:** for every w' in $\text{Dox}(a, w)$, p is true in w'

5. Conclusion

This paper has shown that superordinate negation may target distinct modal components of the attitude verb *craindre*, depending on contextual factors. On the one hand, if the goal of the assertion is belief-related (inquisitive), then superordinate negation will target the foregrounded doxastic component of the verb, and the whole sentence express disbelief. On the other, if the goal of the sentence is dispreference-related (admonitive or reprehensive), then superordinate negation will target the foregrounded dispreference component of the verb – and the whole sentence will express dispreference denial. Based on the observation that expletive negation is infelicitous in the scope of the inquisitive use of *craindre* under negation, I have argued that expletive negation is sensitive to the verb’s modal layering. More specifically, I have argued that *ne* places semantic constraints on the embedding context, requiring it to convey a presupposition of (strong or weak) belief with respect to the prejacent, so that it can express dispreference with respect to the prejacent. This semantic constraint is met when *craindre* receives a volitive reading. I have argued that the presence of expletive negation is correlated to the volitive meaning of *craindre* in the first stages of the history of French, while *ne* later lost its speech-act potential along with the gradual semantic weakening of *craindre* towards a psychological reading.

References

- Ageno, Franca. 1955. L’uso pleonastico della negazione nei primi secoli. *Studi di Filologia Italiana* 13. 339–361.
- Anand, Pranav & Valentine Hacquard. 2013. Epistemics and attitudes. *Semantics and Pragmatics* 6. 1–59. <https://doi.org/10.3765/sp.6.8>
- Chatzopoulou, Katerina. 2012. Negation and Nonveridicality in the history of Greek. Ph.D. Dissertation: University of Chicago.
- Condoravdi, Cleo. 2002. Temporal orientation of modals: Modals for the present and for the past. In David Beaver, Stefan Kaufman, Brady Clark & Luis Casillas (eds.), *The construction of meaning*, 59–88. Stanford, CA: CSLI Publications.
- Dryer, Matthew. 1996. Focus, pragmatic presupposition, and activated propositions. *Journal of Pragmatics* 26. 475–523. [https://doi.org/10.1016/0378-2166\(95\)00059-3](https://doi.org/10.1016/0378-2166(95)00059-3)
- Ducrot, Oswald. 1985. *Le dire et le dit*. Paris: Les Editions de Minuit.
- Espinal, Maria Teresa. 2007. Licensing expletive negation and negative concord in Catalan and Spanish. In Florain Floridic (ed.), *La négation dans les langues romanes*, 47–74. Amsterdam: John Benjamins. <https://doi.org/10.1075/lis.26.04esp>
- Farkas, Donka. 1985. *Intensional descriptions and the Romance subjunctive mood*. New York: Garland.
- Farkas, Donka. 1992. On the semantics of subjunctive complements. In Paul Hirschbühler & Konrad Koerner (eds.), *Romance languages and modern linguistic theory*, 69–105. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.91.07far>

- Gaatone, David. 1971. *Etude descriptive du système de la négation en français contemporain*. Geneva: Droz.
- Giannakidou, Anastasia. 1998. *Polarity sensitivity as (non)veridical dependency*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.23>
- Giannakidou, Anastasia. 1999. Affective dependencies. *Linguistics and Philosophy* 22(11). 367–421. <https://doi.org/10.1023/A:1005492130684>
- Giannakidou, Anastasia & Alda Mari. 2016. Emotive predicates and the subjunctive: A flexible mood OT account based on (non)veridicality. *Proceedings of Sinn und Bedeutung* 19. 288–305.
- Giannakidou, Anastasia & Alda Mari. 2017. A unified analysis of the future as epistemic modality. *Natural Language Linguistic Theory* 36(1). 85–129. <https://doi.org/10.1007/s11049-017-9366-z>
- Giannakidou, Anastasia & Alda Mari. 2020. *Veridicality in grammar and thought: Modality, propositional attitudes and negation*. Chicago: University of Chicago Press.
- Giorgi, Alessandra & Fabio Pianesi. 1997. *Tense and aspect: From semantics to morphosyntax*. Oxford: Oxford University Press.
- Jespersen, Otto. 1917. *Negation in English and other languages*. Copenhagen: A.F.
- Lakoff, Robin. 1968. *Abstract syntax and Latin complementation*. Cambridge, MA: MIT Press.
- Larrivée, Pierre. 1994. Commentaires explétifs à propos d'un certain emploi de ne. *Linguisticae Investigationes* 18(1). 175–186. <https://doi.org/10.1075/li.18.1.09lar>
- Larrivée, Pierre. 2010. The pragmatic motifs of the Jespersen cycle: Default, activation and the history of negation in French. *Lingua* 120. 2240–2258. <https://doi.org/10.1016/j.lingua.2010.03.001>
- Lichtenberk, Frantisek. 1995. Apprehensional epistemics. In Joan Bybee & Suzanne Fleischman (eds.), *Modality in grammar and discourse*, 293–328. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.32.12lic>
- Mari, Alda & Chloé Tahar. 2020. Negative priorities: Evidence from prohibitive and expletive negation. *Proceedings of Sinn und Bedeutung* 24. 56–71.
- Muller, Claude. 1991. *La négation en français*. Geneva: Droz.
- Parry, Mair. 2013. Negation in the history of Italo-Romance. In Anne Breitbarth, Lucas Christopher & David Willis (eds.), *The history of negation in the languages of Europe and the Mediterranean*, Volume I: Case Studies. 59–88. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199602537.003.0003>
- Portner, Paul. 2009. *Modality*. Oxford: Oxford University Press.
- Portner, Paul & Aynat Rubinstein. 2012. Mood and contextual commitment. In *Proceedings of SALT*, 22. 461–487. <https://doi.org/10.3765/salt.v22i0.2642>
- Searle, John. 1969. *Speech acts. An essay in the philosophy of language*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139173438>
- van der Wurff, Wim. 1999. On expletive negation with adversative predicates in the history of English. In Ingrid Tieken-Boon van Ostade, Gunnel Tottie & Wim van der Wurff (eds.), *Negation in the history of English*, 295–328. Berlin: Mouton De Gruyter. <https://doi.org/10.1515/9783110806052.295>
- Villalta, Elizabeth. 2008. Mood and gradability: An investigation of the subjunctive mood in Spanish. *Linguistics and Philosophy* 31(4). 467–552. <https://doi.org/10.1007/s10988-008-9046-x>
- Yanovich, Igor. 2013. Four pieces for modality, context and usage. Ph.D. Dissertation: Massachusetts Institute of Technology.
- Yoon, Suwon. 2011. Not in the mood: The syntax, semantics and pragmatics of evaluative negation. Ph.D. Dissertation: University of Chicago.

Fission in Romance demonstrative-reinforcer constructions

Silvia Terenghi
Utrecht University

This paper proposes a new approach to Romance demonstrative-reinforcer constructions. The account is based on a binary valued feature system for deictic person and is embedded in the Distributed Morphology framework. Looking at data from Romance varieties, some (implicit) shortcomings of previous accounts are repaired via a morphological operation: Fission. Specifically, those accounts do not provide formal means to make sense of the deictic compatibility constraint between the demonstrative and its reinforcer, nor do they discuss categorisation issues relative to reinforcers. Via Fission, instead, a featural reason is given to ensure deictic compatibility, and I put forward a new approach to the category of reinforcers, aiming to overcome their problematic categorisation as DP-internal adverbs.

Keywords: demonstrative-reinforcer constructions, deixis, person, features, Distributed Morphology, Fission

1. Introduction

Romance varieties display demonstrative-reinforcer constructions (henceforth: DRCs; Brugè 1996, 2002; Bernstein 1997, 2001; Roehrs 2010), i.e., combinations of a demonstrative form (either pronominal or adnominal) and a locative element, the reinforcer:¹

- (1) (Italian)
 - a. *questo qua*
this here
 - b. *quel cane là*
that dog there

1. This paper only considers exophoric demonstratives, as the endophoric use of demonstratives has been shown to be secondary with regard to the exophoric one (Diessel 1999: Chapter 5). The extension of the system proposed here to endophoricity seems likely; however, further investigation is needed.

In Italian, the proximal construction (1a) refers to something that is close to the author (or speaker); the distal one (1b) refers instead to something that is far away from the author. These are the only two possible combination patterns for these forms: **questo là* and **quel cane qua* are ungrammatical. The restriction on co-occurrence patterns is referred to as deictic compatibility, that is the deictic features encoded on the demonstrative and on the reinforcer have to be compatible.

In this paper, I address two main questions. The first concerns the derivation of the deictic compatibility restriction on co-occurrence patterns: I propose a post-syntactic approach that derives all and only the attested patterns. A side question is related to the syntactic status of reinforcers, usually regarded as adverbs, in spite of their distributional and combinatorial idiosyncrasies. The analysis I put forward naturally accounts for these facts by considering reinforcers as locative roots (thus, uncategorised elements).

To address these issues, I first describe Romance (mainly Italo-Romance) demonstrative systems and DRCs in featural terms (§ 2). Then I turn to their analysis in terms of Fission, a morphological operation (§ 3). In § 4, I extend the account and present some additional advantages. Finally, I discuss the status of reinforcers (§ 5).

2. Data: Romance demonstratives and DRCs

In this section, I introduce the characterisation of Romance demonstrative systems and DRCs. The following description rests on three core assumptions concerning the featural representation of spatial deixis, for which no in-depth discussion is provided for reasons of space (for details, see Terenghi 2019).

Assumption 1. Spatial deixis is a person-oriented phenomenon, rather than a distance-oriented one. (for a discussion: Diessel 1999: 39ff.)

Deictic features define the location of entities and areas of the external world in relation to a deictic centre. In Romance languages, the deictic centre is most frequently the author, even though some varieties display forms that are centred on both discourse participants (i.e., the author and the hearer), or on the hearer (or addressee) alone (see Table 1 below). Therefore, demonstrative systems in Romance languages can arguably be defined as person-oriented.²

2. In some Romance systems, distance oppositions seem to play a role (e.g., Spanish). In such cases, distance can be conceived as a modifier of the basic person-oriented opposition: see Lander & Haegeman (2018: 51ff.) for degrees of distance as adverbial modifications introduced by an additional structural layer.

Assumption 2. As a person-oriented phenomenon, spatial deixis is best accounted for in terms of person features, rather than of locative ones.

Notice that, even in systems that assume locative features for demonstrative forms, those features are ultimately linked to person. Lander & Haegeman (2018: 3), for instance, define ‘Proximal’ as “close to the speaker”, ‘Medial’ as “close to the hearer” and ‘Distal’ as “far from speaker and hearer”. Therefore, in what follows, spatial deixis will be taken to reduce to person features.

Assumption 3. The person features assumed here are two binary features, [\pm Participant] and [\pm Author]. (see, a.o., Nevins 2007)

Further, deictic person features are different from the person features found within the φ -bundle. The former define proximity relations to one or none of the discourse participants, while the latter are involved in agreement relations. Demonstrative forms always display 3rd person syntax, but make contrastive reference to all available deictic persons. To differentiate deictic person features from φ person features, I add a subscript ‘deixis’ label that restrains the deictic person features to the deictic domain.

Therefore, the deictic features involved in the definition of demonstrative and locative forms are: [\pm Author_{deixis}], for a referent close to (+) or far from (–) the author; [\pm Participant_{deixis}], for a referent close to (+) or far from (–) (one of) the discourse participants (i.e., author and/or hearer).

Table 1 introduces the demonstrative systems attested across the Romance domain. Although they are not discussed here, locative adverbs are comparable, except that there does not seem to be a unary system in this case. Examples come from Ledgeway & Smith 2016 (pp. 890, 879, 886, 885, respectively), where more details on locative adverbs can be found as well.

Table 1. Demonstrative systems

System	Example (Language)*	Deictic person features
Unary	<i>ce</i> (French) DEM	Underspecified
Binary A [\pm Author]	<i>questo / quello</i> (Italian) DEM.PROX / DEM.DIST	<i>questo</i> : [+Author _{deixis}] <i>quello</i> : [–Author _{deixis}]
Binary B [\pm Participant]	<i>chistà / chillà</i> (Neapolitan) DEM.PROX/MED / DEM.DIST	<i>chistà</i> : [+Participant _{deixis}] <i>chillà</i> : [–Participant _{deixis}]
Ternary [\pm Participant, \pm Author]	<i>quiste / quisse / quille</i> (Abruzzese) DEM.PROX / DEM.MED / DEM.DIST	<i>quiste</i> : [+Part. _{deixis} , +Auth. _{deixis}] <i>quisse</i> : [+Part. _{deixis} , –Auth. _{deixis}] <i>quille</i> : [–Part. _{deixis} , –Auth. _{deixis}]

* While I use conventional Leipzig glossing rules, these systems are person-oriented: therefore, ‘PROX[imal]’ encodes reference to the author, ‘MED[ial]’ to the hearer and ‘DIST[al]’ to neither of them.

In unary systems, the single demonstrative form is underspecified for deictic features. Binary systems are optimally defined by just one feature: in some systems, as in Italian, it is [\pm Author_{deixis}], as no reference to the hearer can be encoded. In other systems, as in Neapolitan, the relevant feature is [\pm Participant_{deixis}]: Ledgeway & Smith (2016: 882ff.) refer to the [+Participant_{deixis}] form as “inclusive first person term”. Finally, ternary systems can only be generated by the combination of both binary features.

In DRCs, person-oriented demonstratives combine with person-oriented reinforcers (deictic compatibility being a necessary condition for such a combination, as mentioned in § 1). In Table 2 below, an overview of the possible interactions between the deictic features implied in DRCs is presented (examples from Ledgeway & Smith 2016: 885, 881, 885, 887, respectively).³

Table 2. Demonstrative-reinforcer constructions

Type	Example (Language)	Deictic person features
No co-occurring features	<i>(ë)s(ë) sì / (ë)s(ë) lì / (ë)s(ë) là</i> (Ligurian-Piedmontese) DEM REINF.PROX / DEM REINF.MED / DEM REINF.DIST	DEM: underspecified REINF: [\pm Participant _{deixis} , \pm Author _{deixis}]
Coinciding features	<i>sto qua / queo là</i> (Venetan) DEM.PROX REINF.PROX / DEM.DIST REINF.DIST <i>*sto là / *queo qua</i> DEM.PROX REINF.DIST / DEM.DIST REINF. PROX	DEM: [+Author _{deixis}] \leftrightarrow REINF: [+Author _{deixis}] DEM: [-Author _{deixis}] \leftrightarrow REINF: [-Author _{deixis}]
Partially overlapping features A	<i>cust sì / cul lì / cul là</i> (Piedmontese) DEM.PROX REINF.PROX / DEM.DIST REINF.MED / DEM.DIST REINF.DIST	DEM: [\pm Author _{deixis}] REINF: [\pm Participant _{deixis} , \pm Author _{deixis}]
Partially overlapping features B	<i>chistu ccà / chistu ddhocu / chillu ddhà</i> (Messinese) DEM.PROX/MED REINF.PROX / DEM. PROX/MED REINF.MED / DEM.DIST REINF. DIST	DEM: [\pm Participant _{deixis}] REINF: [\pm Participant _{deixis} , \pm Author _{deixis}]

In the ‘no co-occurring features’ pattern, there is no interplay between deictic features, as they are contrastively expressed on the reinforcer alone. In the ‘coinciding features’ pattern, the deictic features expressed by the demonstrative and the

3. A system split between the ‘no co-occurring features’ and the ‘coinciding features’ patterns is also attested, but it will be left aside here. See Terenghi (2019) for a more comprehensive description of Romance DRCs.

reinforcer are identical, and no other combination of the two forms (and hence feature sets) is possible. Finally, the forms in the two ‘partially overlapping features’ patterns display only partially overlapping deictic features sets, rather than fully identical ones. Such systems are found in varieties that have a binary demonstrative system and a ternary reinforcer system.⁴ In those varieties, the two systems are combined to single out the hearer-related domain (features: [+Participant_{deixis}, –Author_{deixis}]) in marked contexts.⁵

In Piedmontese varieties, the demonstrative binary system is defined by [\pm Author_{deixis}]: *cust* [+Author_{deixis}] and *cul* [–Author_{deixis}] (‘this’ and ‘that’). The reinforcer system instead is ternary, so both features are involved in the featural definition of its forms: *sì* [+Participant_{deixis}, +Author_{deixis}], *lì* [+Participant_{deixis}, –Author_{deixis}] and *là* [–Participant_{deixis}, –Author_{deixis}] (‘here near me’, ‘there near you’, ‘there far away from both’). Therefore, the three Piedmontese DRCs all combine featurally different elements. Notice that the deictic feature encoded on the demonstrative form is a subset of the deictic features encoded by the reinforcer. This explains, in featural terms, why the medial reinforcer *lì* can only be combined with the distal demonstrative, *cul*, and never with the proximal one, *cust*. This system is restricted to the north-western Italian area.

Messinese works along the same lines, but its binary demonstrative system is based on the [\pm Participant_{deixis}] feature: *chistu* [+Participant_{deixis}] and *chillu* [–Participant_{deixis}] (‘this’ and ‘that’). The ternary reinforcer system is defined as follows: *ccà* [+Participant_{deixis}, +Author_{deixis}], *ddhocu* [+Participant_{deixis}, –Author_{deixis}] and *ddhà* [–Participant_{deixis}, –Author_{deixis}]. The deictic feature of demonstrative forms is a subset of the deictic features expressed by reinforcers, which again explains why the medial reinforcer, *ddhocu*, combines with the proximal demonstrative, *chistu*, and not with the distal one, *chillu*. This type is documented for Neapolitan (Ledgeway & Smith 2016: 887) and Brazilian Portuguese (Ledgeway & Smith 2016: 888) as well.

The data overview provided in this section shows that, on a purely descriptive level, the deictic compatibility constraint holding for all attested patterns of DRCs is defined by the feature(s) implied in the characterisation of each item (see Table 2). The question then is how to derive these combinations making sure that no other co-occurrence pattern emerges?

-
4. The reverse case is not attested, as far as I know.
 5. Ledgeway & Smith (2016: 887) and references therein.

3. Fission

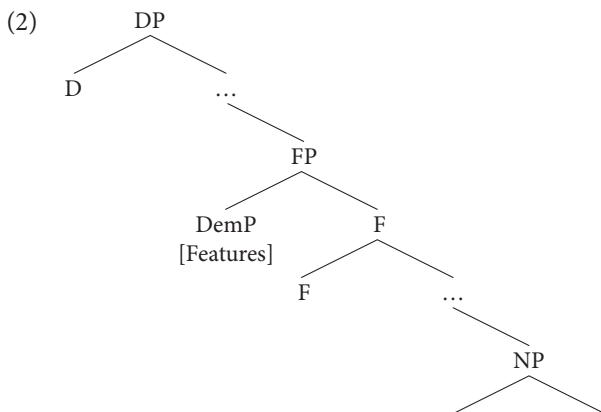
One way forward is to suppose that an Agree relation holds between the two forms. However, this approach implies some featural problems (e.g., stipulation of unvalued features, non-matching feature sets in the ‘partially overlapping features’ pattern; for a complete overview, see Terenghi 2019). Another option would be to assume that hypothetical non-matching features within DRCs get blocked at the interface with semantics because of compositionality issues, within a filtration-based approach to syntax. However, this solution would not be able to naturally account for the seeming dependency of reinforcers on demonstratives: in fact, only demonstratives can occur alone in unmarked contexts, even though their deictic features are just a subset of the reinforcers’ ones.

‘Partially overlapping features’ systems (as Piedmontese and Messinese discussed in the previous section) strongly suggest a different approach: DRCs can be accounted for by resorting to Fission, a morphological operation. The rationale behind the Fission account proposed here is that such co-occurrence patterns can be accounted for by introducing the demonstrative and the reinforcer post-syntactically (in a Distributed Morphology (DM) framework), as concurrent expressions of the same features set. This grants the deictic features a unitary interpretation at the C-I interface, while still making it possible for the actual lexical items to be partly differently specified.

In DM, each Vocabulary Item is conceived as a pair of a phonological representation and of a set of morpho-syntactic features. After narrow syntax, Vocabulary Insertion takes place: for each terminal node, one Vocabulary Item is inserted. Insertion depends on the result of the competition between Vocabulary Items: the most specified one (the one that matches the largest number of features in the terminal node) is inserted and spells out (or discharges) the features of the terminal node. In case the available Vocabulary Items cannot discharge all its features, Vocabulary Insertion ends with underspecification. Fission takes place at this point: the features of the terminal node left underspecified by Vocabulary Insertion can still be exponed, but they will be fissioned, or split off, into a second feature bundle which will be discharged by a different Vocabulary Item. In this way, one single position of exponence in the syntax is split into many in the Morphology. See Noyer (1992) and Halle (1997) for details.

In order to see how a Fission-based account can derive the ‘partially overlapping features’ pattern, it is first necessary to introduce the relevant syntactic structure (2): following standard literature, demonstratives are taken to be phrasal elements (DemP) in the specifier position of a FP. Their featural content (φ -features

and deictic features) is encoded as abstract features under the DemP terminal node and is handed over to the Morphology, where Vocabulary Insertion takes place.^{6,7}



In what follows, I discuss how the derivation works for Messinese (the derivation for Piedmontese follows the same lines, the only difference being the featural composition of the single Vocabulary Items: see § 2). The terminal nodes corresponding to the three available DRCs (*chistu ccà*, *chistu ddhocu*, *chillu ddhà*) are respectively the following ones:

- (3) a. $[\varphi, +\text{Author}_{\text{deixis}}, +\text{Participant}_{\text{deixis}}]$
 b. $[\varphi, -\text{Author}_{\text{deixis}}, +\text{Participant}_{\text{deixis}}]$
 c. $[\varphi, -\text{Author}_{\text{deixis}}, -\text{Participant}_{\text{deixis}}]$

The list of Vocabulary Items for Messinese, based on the discussion in § 2, is provided in (4):

- (4) a. *chistu* $\leftrightarrow [\varphi, +\text{Participant}_{\text{deixis}}]$
 b. *chillu* $\leftrightarrow [\varphi, -\text{Participant}_{\text{deixis}}]$
 c. *ccà* $\leftrightarrow [+ \text{Author}_{\text{deixis}}, +\text{Participant}_{\text{deixis}}]$
 d. *ddhocu* $\leftrightarrow [-\text{Author}_{\text{deixis}}, +\text{Participant}_{\text{deixis}}]$
 e. *ddhà* $\leftrightarrow [-\text{Author}_{\text{deixis}}, -\text{Participant}_{\text{deixis}}]$

6. It can be argued that a definiteness feature linked to the D-level is also involved. However, as it is irrelevant for the analysis, I leave it aside in the following discussion.

7. The syntactic features associated with demonstratives are not relevant for this account. For the sake of consistency, only Vocabulary Items displaying the same φ features are discussed: third person singular masculine (in a tentative featural notation: $[-\text{Participant}_{\text{formal}}, -\text{Author}_{\text{formal}}, +\text{sg}, +\text{masc}]$). I provisionally use the ‘formal’ subscript as the counterpart to the ‘deixis’ one introduced in § 2. ‘Formal’ makes reference to the person features involved in agreement.

As Vocabulary Insertion begins, the form in (4a–e) that is maximally compatible with the features in the terminal nodes in (3a–c) is inserted: that is, either (4a) or (4b), according to the value of the [Participant] feature in the actual terminal node (*chistu* for (3a–b) and *chillu* for (3c)). The demonstrative forms, in fact, discharge all φ features (cf. fn. 7) and part of the deictic features at once, therefore winning the competition against the reinforcers in (4c–e). However, part of the deictic value of the terminal node is left unexpressed: nothing matches the [\pm Author] feature. In unmarked contexts, Vocabulary Insertion ends at this point. In marked contexts, instead, the leftover deictic feature is expressed by second exponence, that is by Fission: the reinforcers provide the missing [\pm Author] values, discharging the terminal node completely.

This derivation suggests that the distinctive features of the binary demonstrative system, i.e., [\pm Participant] in Messinese, are secondary in reinforcers and primary in demonstratives. Secondary features are “properties which must be previously discharged in order for a given rule to apply” (Noyer 1992: 69).⁸ This means that the insertion of reinforcers is conditioned by the featural composition of demonstrative forms, preventing a reinforcer that is specified for a different [Participant] value than that of the demonstrative to be inserted. The features of reinforcers in (4c–e) are accordingly revised, marking secondary features through parentheses:

- (4) c'. $cc\grave{a} \leftrightarrow [+Author_{deixis}, (+Participant_{deixis})]$
 d'. $ddhocu \leftrightarrow [-Author_{deixis}, (+Participant_{deixis})]$
 e'. $ddh\grave{a} \leftrightarrow [-Author_{deixis}, (-Participant_{deixis})]$

Therefore, for the terminal nodes given in (3a–c), we get:

- (5) a. [$\varphi, +Author_{deixis}, +Participant_{deixis}$] \leftrightarrow
 [/'kistu/: $\varphi, +Participant_{deixis}$] + [/'kka/: $+Author_{deixis}, (+Participant_{deixis})$]
 b. [$\varphi, -Author_{deixis}, +Participant_{deixis}$] \leftrightarrow
 [/'kistu/: $\varphi, +Participant_{deixis}$] + [/'d\text{d}ku/: $-Author_{deixis}, (+Participant_{deixis})$]
 c. [$\varphi, -Author_{deixis}, -Participant_{deixis}$] \leftrightarrow
 [/'killu/: $\varphi, -Participant_{deixis}$] + [/'d\text{d}a/: $-Author_{deixis}, (-Participant_{deixis})$]

After φ features and the [Participant] feature have been spelled out by the demonstratives (*chistu* or *chillu*), the conditions imposed by secondary features straightforwardly account for the insertion of *ddh\grave{a}* in (5c) and leave *cc\grave{a}* and *ddhocu* to compete for insertion in (3a) and (3b). Because of their opposed value for [Author],

8. Incidentally, this (along with simple competition) guarantees that demonstratives are inserted as first choice and can stand alone, in case of underspecification, without resorting to any stipulated dependency.

the insertion of *ccà* in (5a) and of *ddhocu* in (5b) is univocal. Any other conceivable combination of demonstratives and reinforcers is naturally ruled out.

In this section, I have shown that agreement problems and dependency stipulations are avoided if DRCs are derived post-syntactically, that is by splitting a single syntactic node (with its own interpretation at the interface with semantics) into two positions of exponence in the Morphology. The demonstrative and the reinforcer are ultimately two complementary exponents for the two scattered feature bundles of DemP, which accounts for deictic compatibility and for co-occurrence restrictions.

4. Beyond deictic fission: [Focus]

The discussion in § 3 was based on one co-occurrence pattern only: in this section, I extend the account to the other co-occurrence patterns introduced in Table 2.

In case of a ‘no co-occurring features’ pattern, Fission applies straightforwardly: the reinforcer discharges the entire deictic value of the terminal node, while the demonstrative discharges its φ component.

The issue is instead more complex in case the involved deictic features coincide, as in (6):

(6) (Italian)

questo qua / quello là
this here that there

In (6), the demonstrative and the reinforcer both encode the same deictic feature: (respectively) either [+Author_{deixis}] or [-Author_{deixis}]. However, on logical grounds, it cannot be maintained that the terminal node includes in its features set the very same feature twice. This makes it impossible to introduce a second exponent, i.e., the reinforcer, as the demonstrative form would discharge the entire featural value of the node itself. Therefore, DRCs would need two different explanations, namely: Fission (‘no co-occurring features’ and ‘partially overlapping features’) and some sort of doubling mechanism (‘coinciding features’).

However, as shown by Terenghi & Casalicchio (2019) for Italo-Romance varieties, locative elements started to combine with demonstrative forms in focal contexts.⁹ This fact suggests that a [Focus] feature is also involved in the terminal node that is ultimately realised as a DRC: under this hypothesis, the relevant

9. Cf. § 2 and Ledgeway & Smith (2016: 887) and references for the marked value of DRCs with ‘partially overlapping features’.

morpho-syntactic features set is [φ , person_{deixis}, Focus].¹⁰ Such a claim is supported by independent evidence provided by Bernstein's (2001) analysis of DP-final elements in Romance languages as focussed elements.

Given the new featural composition of the terminal node, Fission can apply: demonstratives discharge [φ] features and (part of the) deictic features, while reinforcers discharge [Focus]. The choice of reinforcers, once again, is conditioned by the featural composition of demonstratives through secondary features. Vocabulary Insertion for *questo qua* is provided in (7):

- (7) [φ , +Author_{deixis}, Focus] \leftrightarrow [/'kwesto/: φ , +Author_{deixis}] + [/'kwa/: Focus, (+Author_{deixis})]

The addition of a [Focus] feature, beside accounting in a satisfactory way for both the marked interpretation associated to DRCs and for their attested featural patterns, has the further advantage of explaining in a principled way otherwise problematic linearisation issues. Specifically, Romance DRCs display a discontinuous pattern when used in adnominal functions:

- (8) (Italian)
 a. *questo (bel) maglione qui*
 this (nice) sweater here
 'This (nice) sweater here.'
 b. **questo (bel) qui (bel) maglione*

In such cases, the reinforcer cannot possibly be adjacent to the demonstrative, but has to be at the end of the phrase. If we take the reinforcer to be the second exponent of the same terminal node as the one of the demonstrative, it is not clear why the two forms are (and must be) separated. However, given that reinforcers are [Focus]-carrying elements, it can be argued that they move to a phrase-peripheral position to satisfy prosodic well-formedness conditions (see Zubizarreta 1998 for prosodically-driven movement deriving Romance VOS orders). Concretely, as stress is assigned to the rightmost constituent in Italian (Cinque 1993), focussed reinforcers prosodically (i.e., after the Morphology) move to the rightmost edge of the phrase to repair the conflict between the Focus Prominence Rule and the Nuclear Stress Rule.

10. A reviewer asks whether [Focus] is necessary here, or whether the focus reading is prompted by non-syntactic means (specifically, the correlation between extra form and extra meaning). Considering DRCs on a par with the other focussed constructions discussed in Bernstein (2001), it seems worthy to assume [Focus]. Evidence comes from, e.g., Italian possessives, which can be prenominal (Art+Poss+N) or postnominal (Art+N+Poss): only the latter construction gets the focus interpretation, which is however not paralleled by extra lexical items. Hence, I retain [Focus] here, but leave further thinking about the formal and diachronic sides of the issue to future work.

5. The category of reinforcers

Besides accounting for all and only the attested DRCs co-occurrence patterns, the derivation I propose has some welcome consequences for the reinforcers' categorisation, too. Romance reinforcers are identical to locative adverbs (with some exceptions, most notably: French *-ci*, opposed to the proximal adverb *ici* 'here'). Economy considerations suggest that they be regarded as one and the same form, rather than as two homophonous lexical entries. However, it can be argued that reinforcers are not locative adverbs, on both theoretical and empirical considerations.

On theoretical grounds, there should be no such thing as DP-internal adverbs, as adverbs are commonly taken to modify either the verbal or the clausal domain, but never the nominal one. The only apparent exception consists of some specific sets of adverbs that modify process nominals:

- (9) (Greek)
i katastrofi tis polis olosheros
 the destruction the city-GEN completely
 "The destruction of the city completely." (Alexiadou 2001: 47)

Alexiadou (2001: 47ff.) accounts for this use by making reference to the functional structure of process nominals, which includes a domain with verbal properties: those adverbs are linked to the Aspectual/Voice Phrase inside the nominalised verb, rather than to the nominal itself. If reinforcers were to be considered adverbs, then, they would be an unprincipled exception to this generalisation.

Empirical considerations further argue against a unitary account for locative adverbs and reinforcers. In fact, it can be observed that reinforcers display some idiosyncrasies that distinguish them from locative adverbs.

First, consider adverbial modification:

- (10) (Italian)
 a. *laggiù / solo qui*
 there-down only here
 b. *#quel cane laggiù*
 that dog there-down
 c. **questo cane solo qui*
 this dog only here

While locative adverbs can be modified (10a), this is not the case for reinforcers, for which modification by other locative forms yields semantically deviant phrases (10b), whereas modification by focalising adverbs results in ungrammaticality (10c).¹¹

11. Notice that, although French *là-bas* 'there' [*lit.* 'there-down'] is an apparent exception, *-ci* 'here' and *-là* 'there' still cannot be modified by focalising adverbs (e.g., *seulement* 'only'): **ce chien seulement-ci* 'this dog only-here'.

Secondly, coordination tells apart reinforcers and locative adverbs. For deictic compatibility reasons, disjunctive coordination of *lì* and *là* (respectively: specific/punctual and generic/areal ‘there’) is taken into consideration here:

(11) (Italian)

- a. *ti* *siedi* *lì* *o là?*
 you.REFL sit.2SG.PRS there.SPECIFIC or there.GENERIC
 “Are you going to sit there or (over) there?”
- b. **quel* *ragazzo lì* *o là?*
 DEM.DIST boy there.SPECIFIC or there.GENERIC
 “That boy there or (over) there?”

Contrary to locative adverbs (11a), reinforcers cannot be coordinated by disjunctive *o* ‘or’: (11b) cannot possibly make reference to two different boys located in different places (as it would be the case for *quel ragazzo lì o quel ragazzo là?* ‘that boy there or that boy (over) there?’), nor of course can it refer to one and the same boy.¹²

Under the Fission account proposed here, reinforcers are additional exponents of DemP, inserted post-syntactically and conditioned by the deictic feature(s) encoded by the demonstrative. Instead, no reference is made to their syntactic environment. This suggests that reinforcers be treated as locative roots that cannot possibly be in any suitable c-commanding configuration so as to assume a category, and that they cannot project. Uncategorised locative roots eventually become locative adverbs when merged in the suitable environment. This way, differences in behaviour between reinforcers and locative adverbs are derived, without overloading the Lexicon.

6. Concluding remarks

The core idea of this paper is that Fission can derive the attested combinations of demonstratives and reinforcers in Romance languages, while also ruling out the non-attested patterns. This proposal has two important advantages: first, it naturally accounts for deictic compatibility in a formal way, i.e., considering the featural composition of the combined elements. This makes sense of patterns of co-occurrence of apparently incompatible features, too. Second, it introduces a new analysis for

12. As a reviewer points out, the contrast in (11) could rather be related to the syntactic difference between verbs and nouns. Namely, the verb may have ATB moved out of the coordinated constituent (11a), while such a movement is not available for nouns (11b). More investigation is needed, but for now it can be highlighted that the expression *qua e là* ‘here and there’ is idiomatic, e.g., in *andare qua e là* ‘to wander’ (and crucially not, literally, ‘to go here and (to go) there’).

reinforcers: no longer DP-internal adverbs, but simple non-projecting exponents of a fissioned feature bundle.

In addition to the core analysis, I highlighted some open issues, such as these: the possible role that the [Focus] feature plays within Romance demonstrative systems (is just deictic Fission possible, without [Focus] being involved? can [Focus] be the trigger for Fission?) and linearisation issues (does Fission have an intrinsic directionality?). They are left here as open questions, but hopefully they will help to further understand the nature and interaction of deictic features as encoded by demonstrative forms.

Acknowledgements

I would like to thank my supervisors Roberta D'Alessandro and Jan Casalicchio, the audience of Going Romance 2018, and two anonymous reviewers for their extremely insightful and inspiring comments. All errors are of course my own.

Funding

This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement: CoG 681959_MicroContact).

References

- Alexiadou, Artemis. 2001. *Functional structure in nominals: Nominalization and ergativity*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.42>
- Bernstein, Judy B. 1997. Demonstratives and reinforcers in Romance and Germanic languages. *Lingua* 102. 87–113. [https://doi.org/10.1016/S0024-3841\(96\)00046-0](https://doi.org/10.1016/S0024-3841(96)00046-0)
- Bernstein, Judy B. 2001. Focusing the “right” way in Romance Determiner Phrase. *Probus* 13. 1–29. <https://doi.org/10.1515/prbs.13.1.1>
- Brugè, Laura. 1996. Demonstrative movement in Spanish: A comparative approach. *University of Venice Working Papers in Linguistics* 6. 1–61.
- Brugè, Laura. 2002. The positions of demonstratives in the extended nominal projection. In Guglielmo Cinque (ed.), *Functional structure in DP and IP: The cartography of syntactic structures*, 15–53. Oxford: Oxford University Press.
- Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24. 239–297.
- Diessel, Holger. 1999. *Demonstratives. Form, function, and grammaticalization*. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.42>

- Halle, Morris. 1997. Distributed Morphology: Impoverishment and fission. In Benjamin Bruening, Yoonjung Kang & Martha McGinnis (eds.), *PF: Papers at the Interface* [MITWPL 30], 425–449. Cambridge, MA: MIT Press.
- Lander, Eric & Liliane Haegeman. 2018. The nanosyntax of spatial deixis. *Studia Linguistica* 72(2). 362–427. <https://doi.org/10.1111/stul.12061>
- Ledgeway, Adam & John C. Smith. 2016. Deixis. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford guide to the Romance languages*, 879–896. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199677108.003.0054>
- Nevins, Andrew. 2007. The representation of third person and its consequences for person-case effects. *Natural Language & Linguistic Theory* 25. 273–313. <https://doi.org/10.1007/s11049-006-9017-2>
- Noyer, R. Rolf. 1992. Features, positions and affixes in autonomous morphological structure. Ph.D. Dissertation: Massachusetts Institute of Technology.
- Roehrs, Dorian. 2010. Demonstrative-reinforcer constructions. *Journal of Comparative Germanic Linguistics* 13. 225–268. <https://doi.org/10.1007/s10828-010-9038-4>
- Terenghi, Silvia. 2019. Demonstrative-reinforcer constructions and the syntactic role of deictic features. *Linguistics in the Netherlands* 36. 192–207. <https://doi.org/10.1075/avt.00032.ter>
- Terenghi, Silvia & Jan Casalicchio. 2019. Demonstrative-reinforcer constructions: Variation in Italo-Romance. Paper presented at 14th Cambridge Italian Dialect Syntax-Morphology Meeting (CIDSM 14), Prague, May 27–29.
- Zubizarreta, María L. 1998. *Prosody, focus, and word order*. Cambridge, MA: MIT Press.

Index

A

Acceptability Judgment Task (AJT) 25, 29, 38–43, 197, 199–200
Adjective Phrase (AdjP) 249–250, 253–255, 259
Adverbial Perfect Participial clauses (APC) 264–276, 281–283
Agreement 14, 17, 20, 29, 30, 32, 36n10, 40, 83, 151, 154–155, 162–167, 171–173, 175–185, 187, 225–227, 238–240, 244, 305, 309n7, 311
Agree 14, 159, 161, 163–166, 308
Rich agreement 173, 181, 187
Subject-verb agreement 29, 30, 32, 36n10, 40, 172, 177, 179
Aggregate (Aggr) 159, 167n9
Albanian 157, 288
Italo-Albanian 167
Anaphora 11–12, 14, 15, 17, 71–72, 77–80, 84n6, 86–88, 93–94
Animacy 117–118, 123–128
Antecedent 11–22, 77–78, 80, 83, 87–94, 98, 108, 172
Anteriority 83, 264, 267, 270–278, 280–282
Aspect
Perfective aspect 30, 72–82, 84, 86–87, 89–90, 93–94, 231, 266–267, 274–281
Imperfective aspect 72, 74–78, 81–87, 90–94, 177, 231n7, 238–239, 243, 274, 283
Progressive aspect 26, 33–34, 37, 44–45, 73, 75–77, 81, 93, 225–227, 229–234, 237–241, 244

Prospective aspect 73, 75, 231, 240–241
Viewpoint aspect 71–73, 78
Aspectual types 270
Culminated Process (CP) 270, 272n11, 276–282
Process (Pro) 270, 272n11, 276–281
Culmination (Culm) 270, 276–283
Point (Pon) 270
State (St) 86, 93, 231, 236n11, 270, 272, 276–281, 283
Assertion, (time of) 73, 81, 231, 290–293, 298–301
Assertion purpose 291
Inquisitive 291–292, 296, 299, 301
Reprehensive 291, 295–296, 300
Admonitive 291, 293, 296, 298, 301
Warning 293
Reproach 295–296, 300
Apology 295, 300
Author 304–307, 309–312

B

Basque 100–101, 142–143, 147
Binding (theory) 11, 13, 15–16, 71–72, 77–78, 80–81, 83–88, 91, 93, 109, 111–112, 180n7
Bilingualism 65, 131, 133, 145
Binding Principles
Principle A 12, 15–16
Principle B 11–14, 16–17, 21–22
Principle C 21
Borrowing 131, 138, 143–145, 147

C

Capeverdean 225–228, 230, 232, 235–236n11, 238–239, 241, 243
Case
accusative case 26–30, 32–34, 37–45, 120, 124–126, 140–142, 146, 162
dative case 28–29, 37, 42–44, 140–141, 146, 163
nominative case 26–30, 33, 36, 38–39, 41–44, 124, 176, 178, 181, 183–184, 188
Catalan 55, 58, 100–101, 113, 117–121, 140–142, 147, 205–208, 210–211, 215, 217–218, 221, 250
Eastern Catalan 207
Old Catalan (OC) 117–125, 127, 130, 250
Causative 13, 25–27, 29–31, 33, 38, 42, 45, 137
Clause
Main Clause (MC) 19–20, 111–113, 235, 264, 266–268, 270, 272–283
Subordinate Clause (SC) 264, 268, 270, 273, 275–282, 288
Code switching 131, 133, 137–138, 143–145, 147
Complementizer Phrase (CP) 30, 34, 36, 50, 73, 98, 105–108, 110–113, 156, 160–164, 249, 251, 255–257, 259
Confirmation Phrase 102
Contrastive Left-Dislocation (CLD) 110
Co-phonologies 220
Copula 192, 194, 199–201, 203, 239n14

Coreference 13, 71–72, 77–79,
84, 86–88, 92–94
Culminated event 26, 33 (see
also *Culminated processes*,
under *Aspectual types*)

D

Definiteness 118, 309n6
Degree
Degree word 250–251
Degree Phrase (DegP)
254–255
Deixis 303–307, 309–312
Delay of Principle B 12
Demonstrative 151, 167n9,
303–315
Demonstrative-Reinforcer
Construction (DRC) 303,
305–309, 311–313
Determiner(s) 142, 151, 153, 156,
158–159, 250–252, 253–254n4
D-feature 172, 188
Differential Object Marking
(DOM) 25, 28–29, 33, 37,
40, 43–45, 117–125, 127
Differential Plural Marking
(DPM) 151, 160, 167
Distributed Morphology 131,
135, 303, 308
Dynamicity 191, 196, 199–203

E

Echo-questions 52, 55
Ellipsis 71–72, 77, 87–90,
93–94, 97–98, 103–111, 113–
114, 171, 175, 182, 187
English 11–12, 16, 26, 34n9, 49,
53–54n2, 76–77, 89n9, 134,
136, 138–139, 143, 145, 147,
172, 185n9, 208, 217, 226, 230,
234, 251–252, 266, 268–269,
272n11, 276, 282
American English 109
British English (BE) 266,
268–278, 280–282
Epistemic (verbs, modals,
meaning) 30, 235–237, 290,
298–300

Exceptional Case Marking
(ECM) 13, 16–22, 25–34,
36–38, 41–45
Externalization parameter
160–161

F

Feature Reassembly Hypothesis
(FRH) 27, 35–37, 45
Fission 303–304, 308, 310–312,
314–315
Focus 49–50, 52–53, 55, 62–66,
77, 98–103, 106, 108, 110–111,
256–257, 259, 311–312, 315
Focus fronting (FF) 49–52,
55, 62, 65
Contrastive focus 106
Focus Phrase (FocP) 50, 64,
100–101, 257
French 71–72, 75, 77, 82n3,
89n9, 143, 145, 217, 250,
255–257, 287–288, 293–296,
298–299, 301, 305, 313n11
Middle French 255n5, 295
Modern French 295, 299
Old French 252, 255n5, 293,
295, 297–298
Preclassical French
295–296, 298
Friulian 153
Full Access Hypothesis 192
Full Transfer 192, 196

G

Gender 14–15, 18, 20, 60n9,
131, 136, 138, 140, 142–143, 147,
157–159, 164–166, 171, 175,
182–183, 185
German 64n11, 112, 137–138,
145, 147, 254
Germanic 110
Giküyü (see *Kikuyu*)
Greek 53, 66, 158, 240, 288, 313

H

Harmonic Grammar 205, 207,
215, 220–221
Hindi 112–113
Hungarian 53, 66, 112

I

I-language 131, 133–134, 139,
142–147
Illocutionary denegation 290
Impersonal 36 (footnote),
163, 172
Inflection(s) (Infl) 151–155,
157–159, 162–165, 171, 176–178,
183, 187, 251
Infinitive 17, 19–20, 31–34,
36n10, 38–39, 41, 43–45, 137,
155
Inflected Infinitive Structure
(IIS) 26–27, 29–33, 36–45
Prepositional Infinitival
Structure (PIC) 26–27,
29, 33–34, 36–41, 43–45
Interrogative 49–51, 54–55, 63,
99–101, 104, 111–112
Split Interrogative (SI)
97–99, 104–105, 109, 113
Matching Split Interrogatives
(MIS) 98–101, 103,
109–110
Non-matching Split
Interrogatives (NMSI)
97, 99–114
Italian 11–12, 16, 18, 52, 54–55,
61, 63–65, 151–153, 157–159,
164, 191–196, 198–203, 249–
250, 303–307, 311–314
Old Italian (OI) 249–253,
255–260
Southern Italian 192–194,
199
Italo-Romance 250, 304, 311

K

Kikuyu 241

L

Ladin 153, 156, 158, 160–161,
165–166
Latin 127, 152 (footnote),
249–251, 288, 297–298
Left Branch Extraction (LBE)
251–252
Left periphery 50–51, 99–101,
109, 112, 251, 256–257, 259

- (Ligurian-) Piedmontese 306
 Loanwords 205–212, 216–217, 221
 Locality 12–13, 15, 21–22, 180n7
 Locative 191–192, 194–197, 199n2, 202–203, 239–240, 244, 303–305, 311, 313–314
 Logical Form (LF) 132–133, 140
- M**
 Markedness (constraints) 205–206, 216–218, 221
 Medial reinforcer 307
 Messinese 306–310
 Modal (verbs) 17, 19, 34, 89n9, 225–227, 229–230, 235–238, 242–243, 289, 301
 Modifier 250, 254, 260, 304n2
 Mood 136, 238–241, 243–244, 268, 293–295
 Indicative 177–178, 293–294
 Subjunctive 199, 238, 294–295
- N**
 Neapolitan 198, 305–307
 Negation (marker) 28, 30–31, 35, 101, 107, 110, 135, 287–290, 292, 294, 297, 301
 Expletive negation 287–291, 294–295, 298–299, 301
 Negative evidence 196
 Negative imperative 297
 Nominal class 151, 157, 159
 North Lombard 151, 153, 167
 Nuclear pitch accent (NPA) 49–61, 63, 65–66
 Nuclear stress 53, 56, 312
 Null subject 171–178, 181–185, 187, 249 (see also *Pro (drop)*)
 Number 26, 29, 32, 135, 140, 151, 157–159, 164–166, 171–172, 175–180, 182–183, 185, 187
 Number feature 32, 37, 166, 172, 179, 183
 Number agreement 166, 182
- P**
 Parser 14–15, 17, 20–22
 Participant 18–19, 21, 38, 40, 44–45, 164, 304, 307, 309–310
 Perception verbs 17, 25–26, 29–31, 33, 37, 41–42, 44–45
 Person 35–36
 Phase 30, 151, 159–164, 167, 249–251, 259–260, 270n9
 Phase Impenetrability Condition (PIC) 160–161
 Phonetic Form (PF) 98, 108, 131–133, 139–140
 Phi-features (ϕ -features) 14, 160–164, 166–167, 308
 Picture-naming production test 211–212
 Picture Selection Task 197–198, 201, 203
 Piedmontese 306–308
 Portuguese 226–227, 233–234, 238–240, 243–244, 266, 269
 Angolan Portuguese 266, 272n11
 Brazilian Portuguese (BP) 153, 156, 159–160, 166, 171, 266, 307
 European Portuguese (EP) 25, 27, 35, 42n13, 171, 263–264, 266
 Mozambican Portuguese 266
 Posteriority 264, 266–267, 270–271, 273–279, 282
 Presupposition 50, 290–292, 294–295, 297–301
 Principle of Recoverability of Deletions 14
 Pro (drop) 33, 163, 171–175, 182, 185n9, 187
 PRO 33, 185n9
 Pronominal clitic(s) 11–22, 28–30, 101, 120, 124, 131, 133, 140–142, 147, 151, 153–156, 162–164, 167, 249, 256
 Clitic climbing 11, 17, 28
 Clitic doubling 101, 120, 124
 Clitic trace 17
 Object clitic (OCl) 17, 154–155, 162–163n4
 Reflexive clitic 18 (see also *Reflexive (clitic, pronoun)*)
 Subject Clitic (SCL) 153, 155
 Proximal 304, 307, 313
- R**
 Reconstruction 105–109, 111, 113
 Reflexive (pronoun, clitic) 12–13, 15, 18, 28–29, 122, 163
 Request 290, 298
 Revised Hierarchical Model 133
 Rhaeto-Romance 153
 Romanesco 193–194, 198–199, 202–203
 Romanian 53–56, 66, 120, 122, 124
 Old Romanian 118, 120, 122
- S**
 Sardinian 49–52, 55, 57–65, 259
 Logudorese Sardinian 57
 Scope Marker Constructions (SMC) 112–114
 Scrambling 251, 253–254n4, 258, 260
 Separationism 131–132, 145
 Simultaneity 82, 235–237, 270, 272, 274, 276–280, 282
 Small Clause 30, 33–34, 37, 101, 253, 255–256
 Spanish 25–31, 36–37, 40–46, 55, 58, 71–72, 75, 77, 82, 84, 89n9, 97, 99, 101–107, 109, 111, 113, 119, 136–138, 140–143, 147, 157, 166, 191–200, 202–203, 208, 217, 222, 304n2
 Old Spanish 119
 Specificity 33, 117–118
 Speech act 103, 143, 288, 297–301
 Spell Out 125, 137–140, 146, 161, 163n4, 250, 288
 Strong Cross Over Effect 21
 Strong pronoun 12–13, 16, 21–22, 121

- Swahili 136, 138–139, 147
 Syntactic transfer 131, 145, 147
- T**
- Temporal anaphora 72, 77–79, 88, 93–94
 Temporal binding 72, 80
 Temporal boundedness 191
 Temporal coreference 72, 79
 Temporal Remoteness
 Morphemes (TRMs) 241
 Tense 71–75, 81–84, 86, 93, 177, 225–227, 229, 234–235, 237, 240, 242–244, 267–268, 270, 273–276, 278–282
 Sequence of Tense (SoT) 71–72, 81–83, 86–89, 93–94
 Tense Phrase (TP) 30, 72, 74–75, 80, 89, 92, 98, 108, 136, 163–165, 175n4, 255–257
- Time**
- Assertion Time (AST-T) see (*Assertion, (time of)*)
 Event Time (E) / (EV-T) 73, 75, 77–81, 83–86, 89–93, 231n7, 241
 Speech Time (S) 231n7
 Situation Time (T-Sit) 231–232, 242–243
 Reference Time (R) / (Ref-T) 74–75, 77, 79–81, 84–86, 231n7
 Time of Utterance (TU) 231–234, 237, 239–244
 Topic Time (TT) 231–234, 237, 239–244
- Translanguaging 147
- V**
- Voice 135, 137, 162–165, 313
 Vowel laxing 207, 217
- Vowel reduction 205–207, 211, 215, 217–218, 221
 vP 21, 30, 34, 72, 136–137, 146, 156, 160–164, 175, 256–257, 259
- W**
- Weighted Scalar Constraints 205, 207, 215–216, 219
 Wh-movement 49–53, 55, 62, 65, 102
 Wh-questions 49–52, 54–60, 62–66, 97–98, 108, 113
- Z**
- Zeit-P 72–73, 75, 79–80

This volume contains a peer reviewed selection of invited contributions, papers and posters that were presented at the 2018 venue of Going Romance (XXXII) in Utrecht (a four day program that included two thematic workshops).

The papers all discuss data and formalized analyses of one or more Romance languages or dialects, in either synchronic or diachronic perspective, and pay particular attention to the variation and the actual variability that is at stake, not only in syntax and morpho-syntax but also in semantics and phonology. Beyond the discussion of differences between languages and/or dialects from a formalist perspective, the volume also contains a number of papers linking the theme of variation to sociolinguistic issues such as natural bilingualism and micro-contact.

ISBN 978 90 272 1012 8



9 789027 210128

JOHN BENJAMINS PUBLISHING COMPANY