

Romance Languages and Linguistic Theory 2017

EDITED BY
Alexandru Nicolae
Adina Dragomirescu

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 355

Alexandru Nicolae and Adina Dragomirescu (eds.)

Romance Languages and Linguistic Theory 2017
Selected papers from 'Going Romance' 31, Bucharest

ROMANCE LANGUAGES AND LINGUISTIC THEORY 2017

SELECTED PAPERS FROM
'GOING ROMANCE' 31, BUCHAREST

Edited by

ALEXANDRU NICOLAE

ADINA DRAGOMIRESCU

*'Iorgu Iordan – Alexandru Rosetti' Institute of Linguistics
& University of Bucharest*

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.355

Cataloging-in-Publication Data available from Library of Congress:
LCCN 2021042000 (PRINT) / 2021042001 (E-BOOK)

ISBN 978 90 272 1005 0 (HB)

ISBN 978 90 272 5842 7 (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 <i>Adina Dragomirescu and Alexandru Nicolae</i>	1
The acquisition of verbal passives by Portuguese-speaking children: Some data from comprehension <i>Celina Agostinho and Anna Gavarró Algueró</i>	9
<i>Plus</i> in the French negative system: A presuppositional and non-quantificational n-word <i>Pascal Amsili and Claire Beyssade</i>	29
An experimental approach to parallelism in ellipsis: The case of <i>pro</i> -drop in Romance gapping <i>Gabriela Bîlbîie and Israel de la Fuente</i>	49
On focal and <i>wh</i> -projections, indirect <i>wh</i> -questions, and quantificational chains <i>Caterina Bonan</i>	73
Is there a dative alternation in Romanian? Remarks on the cross-categorical variation of datives in ditransitive constructions <i>Alexandra Cornilescu and Alina Tigău</i>	91
The interpretation of null subjects in Romanian: An information-structure approach for comparative analysis <i>Mara Frascarelli</i>	111
Verum focus and Romanian polar questions <i>Ion Giurgea</i>	135
The downward grammaticalisation of irrealis subordinators in Romanian, Salentino and southern Calabrese <i>Kim A. Groothuis</i>	157
Differential object marking: What type of licensing? <i>Monica Alexandrina Irimia</i>	171

The effects of language ecology on syntactic structure: A look at Kristang and Makista <i>Robert W. Laub</i>	193
The syntactic distribution of <i>raddoppiamento fonosintattico</i> in Cosentino: A phase-theoretic account <i>Adam Ledgeway</i>	205
The causative-inchoative alternation (as we know it) might fall short: Crosslanguage systematicities and untapped data from Romance and Greek <i>M. Eugenia Mangialavori Rasia</i>	239
On <i>wh</i> -extraction in <i>de+que</i> constructions in Spanish <i>Gabriel Martínez Vera</i>	263
On another apparent violation of the subject-island constraint in French <i>Guido Mensching and Franziska Werner</i>	277
Moving towards an event: The Romanian prepositional supine construction <i>Elena Soare</i>	297
Cyclicity without containment in Romanian perfects <i>Donca Steriade</i>	311
Dative clitics in Romanian ditransitives <i>Alina Tigău and Klaus von Heusinger</i>	335
Syntactic vs pragmatic passive: Evidence from Romanian <i>Andra Vasilescu</i>	357
Index	373

Introduction

Adina Dragomirescu and Alexandru Nicolae

'Torgu Jordan – Alexandru Rosetti' Institute of Linguistics /
University of Bucharest

This volume in the series *Romance Languages and Linguistic Theory: Selected papers from 'Going Romance'* contains a selection of papers presented at the 31st Going Romance conference organized in Bucharest on December 7–9, 2017; a special workshop titled 'Substratum and adstratum in the development of Romance morphosyntax' was also held on the last day of the conference.

Contents of the volume

Going Romance is a now established European conference series focused on Romance languages from the perspective of current linguistic theorizing. In its three decades, it has developed into a major European discussion forum for theoretically relevant research on Romance languages where current ideas about language in general and Romance languages in particular are discussed. The present volume does justice to this tradition.

The volume contains 18 double anonymous peer-reviewed papers, including those presented by the keynote speakers (Donca Steriade and Adam Ledgeway). Phenomena found in major Romance languages – European Portuguese (Agostinho & Gavarró Algueró), French (Amsili & Beyssade; Mensching & Werner), Italian (Bonan; Frascarelli), Spanish (Bilbiüe & de la Fuente; Martínez Vera), Romanian (Cornilescu & Tigău; Frascarelli; Giurgea; Groothuis; Soare; Steriade; Tigău & von Heusinger; Vasilescu) – and in dialects – Cosentino and an overview of the southern Italian dialects (Ledgeway), Salentino and southern Calabrese (Groothuis), Trevigiano (Bonan) and Neapolitan (Irimia) – either benefit from in-depth analyses confined to one single variety (Agostinho & Gavarró Algueró; Amsili & Beyssade; Cornilescu & Tigău; Frascarelli; Giurgea; Martínez Vera; Mensching & Werner; Soare; Steriade; Tigău & von Heusinger; Vasilescu) or provide comparative analysis: dialect vs standard language (Bonan), dialect vs different major language(s)

(Groothuis; Irimia), cross-dialectal comparison (Ledgeway), cross-Romance comparison (Bilbîie & de la Fuente) and even comparison of language families (Mangialavori Rasia). Creoles have not been neglected either (cf. Laub's contribution on Makista and Kristang). Theoretical and experimental approaches complement one another, as do diachrony and synchrony. Although contributions to syntactic theory and analysis are dominant, other fields are also represented: phonology and morphophonology (Steriade), acquisition (Agostinho & Gavarró Algueró) and the interfaces (morphophonology-syntax: Ledgeway; syntax-semantics: Amsili & Beyssade; syntax-intonation: Giurgea; syntax-information structure: several chapters).

Celina Agostinho & Anna Gavarró Algueró's study focuses on the acquisition of passives, and addresses the issue of whether verbal passives are delayed in language acquisition. Syntactic homophony (cf. Babyonyshev et al. 2001) does not hold in European Portuguese, as this is a language which distinguishes verbal and adjectival participles by means of different auxiliaries, thus forming a fertile ground for testing hypotheses related to the acquisition of passive constructions. The results reported in Agostinho & Gavarró Algueró's study show that European Portuguese-speaking children display contrasts in the acquisition of actional vs perception verbs and have difficulties with passives of (some) subject experiencer verbs until school age.

Pascal Amsili & Claire Beyssade set out to determine the place of *plus* "no more" in the French system of negation by addressing the following question: is *plus* the presuppositional counterpart of *pas* ("not") or does *plus* belong to the class of n-words? By applying a robust set of distributional and interpretative diagnostics, the authors show that *plus* comes in two guises: (i) *plus* incorporates a sentential negation, and (ii) *plus* has only a presuppositional contribution to the sentence. Furthermore, depending on its position, *plus* adduces different presuppositional contributions: preverbal *plus* prefers an existential presupposition, while postverbal *plus* prefers the universal presupposition.

Gabriela Bilbîie & Israel de la Fuente's analysis of parallelism constraints in gapping in Romance *pro*-drop languages (Romanian and Spanish) reveals that syntactic parallelism is less strict than traditionally assumed. Bilbîie & de la Fuente show that syntactic parallelism is not established at the level of phrase structure, but rather at the level of grammatical functions as listed in the argument structure of predicates. Furthermore, discourse and information structure constraints are stronger than syntactic ones. The analysis of *pro*-drop languages also brings evidence for a topic-focus pattern – rather than a double focus pattern – for gapping in general and for gapping with *pro*-drop in particular.

Caterina Bonan examines the cartography of focus on the basis of novel data from non-standard Italian and a Venetan dialect, Trevigiano, varieties in which focus is less constrained than in standard Italian. Bonan shows that in these varieties, all chain-formation problems are bypassed by particular syntactic properties such as the availability of clitic doubling for foci and the direct left-peripheral merger of adverbials, as well as by the fact that their low (*vP/VP*) peripheral area may host *wh*-phrases.

Alexandra Cornilescu & Alina Tigău set out to demonstrate the existence of a genuine dative alternation in Romanian by examining the internal structure of both inflectional and prepositional datives. They show that both dative forms require a dual categorial analysis (manifesting either DP or PP properties) even in Recipient constructions, which allows for the licensing of their case and person features. The sensitivity of datives to the Animacy Hierarchy indicates the presence of a syntactic [Person] feature, which triggers the reanalysis of the dative preposition as a D-category, the D head thus acting as a category shifter from PP to DP.

By proposing an analysis which shifts the focus from the licensing to the interpretation of null subjects, **Mara Frascarelli** explores the acceptability of referential null subjects in Romanian (against different syntactic conditions and distinct clause types), and shows that Romanian is a consistent *pro*-drop language which shows none of the features that occur in partial null subject languages. However, closer comparison with Italian (also a consistent *pro*-drop language) reveals interesting contrasts: (i) Romanian shows a preference for the closest antecedent possible, which is the effect of a locality requirement; (ii) in Romanian, the usage of overt pronouns triggers a focus interpretation, possibly related to a corrective effect.

Ion Giurgea analyses the stress patterns of polar questions in Romanian, a language in which only intonation systematically distinguishes polar questions from declaratives. Against most of the previous literature on the intonation of Romanian polar questions, which claims that the neutral pattern has the nuclear accent on the finite verb and a final contour characteristic of questions with early focus, Giurgea argues that, at least in information-seeking questions, this pattern is not neutral, but, rather, is an indication of ‘verum focus’, nuclear stress on the last prosodic word actually constituting the neutral pattern. More generally, Giurgea’s chapter contributes to understanding the conditions of the use of verum focus in questions.

Kim A. Groothuis proposes a diachronic analysis of several complementizers from a number of Romance varieties characterized by a more limited use of infinitival complementation (selecting instead a finite clause for irrealis complements): Salentino *cu*, southern Calabrese *mu* (with its variants *ma* and *mi*) and Romanian *să*. By tracing their etymologies (and putting forward a new etymon for *cu* and *mu*, namely Lat. QUOMODO ‘how’) and applying several synchronic diagnostics,

Groothuis analyses all these complementizers as instances of downward (secondary) grammaticalization, arguing that downwards grammaticalization is allowed since these are all functional elements which assume a novel functional role (one interesting speculative consequence of this analysis is that upwards movement in grammaticalization is obligatory only when the grammaticalizing element is a 'lexical' element).

Building on recent work according to which differential object marking (DOM) signals nominals that must undergo licensing in the clausal syntax as they bear an uninterpretable case [uC] feature, **Monica Alexandrina Irimia** brings into the discussion data from Romanian and Neapolitan which indicate that an additional licensing mechanism is necessary in the derivation of DOM. More precisely, some classes of nominals (which include DOM animates under certain conditions) contain a [+PERSON] feature which needs to be valued/checked, and the prepositional DOM marker is interpreted as tracking this additional licensing operation on objects that might have already valued/checked their [uC] feature.

Robert W. Laub investigates the grammar of two related – generally mutually intelligible, yet distinct – Portuguese-lexified creoles, Kristang and Makista. Focusing on two phenomena (genitive patterns and cleft constructions), Laub explores the differences (rather than the commonalities) between these two creoles, and accounts for them from the perspective of language ecology (i.e., the hypothesis that environments in which languages are employed affect their development). The qualitative and quantitative methods employed suggest that there is a connection between language ecology and morphosyntactic structures.

Adam Ledgeway's chapter, presented as a plenary talk, makes important contributions to phase theory and to the understanding of the syntax-phonology interface. Ledgeway starts with an overview of the distribution of *raddoppiamento fonosintattico* (RF) "phonosyntactic doubling" (a phonological fortition process) in the Calabrian dialect of Cosenza, and argues for interpreting locality more broadly, in terms of phasal domains. In contrast to other southern Italian varieties (e.g., Neapolitan), in Cosentino RF has a broader application and occurs if Word₁ (the RF-trigger) and Word₂ (the word undergoing RF) are in the same phasal domain. What is more, certain interpretative phenomena with RF correlates (e.g., in the dialect of Cosenza RF helps to formally distinguish between postverbal definite referential DPs, be they topical or focal, and their non-referential variants; also, RF on the complementizer ensures the availability of coreference between the matrix and the embedded subject) are best understood from the perspective of the various structural constraints in the mapping from syntax to information structure. Ledgeway then extends the analysis to an impressive number of varieties showing RF and, by adopting D'Alessandro & Scheer's (2015) Modular Phase

Impenetrability Condition, shows that it is not always possible to postulate a direct mapping of syntactic Spell-Out domains onto phonological domains.

M. Eugenia Mangialavori Rasia takes issue with the claim that internal arguments are stable/constant arguments in the causative-inchoative alternation and shows that an external-argument-only variant is also available in Romance (and Greek). The derivation of eventless, external-argument-only (monadic) causatives is based on direct composition with a causative external-argument introducing v^0 not complemented by an internal-argument-introducing verbal head.

By assuming Bošković's (2015, 2016) 'Phase Collapsing' theory and 'The Phase-over-Phase Constraint' and Nunes's (2014, 2016) approach to successive cyclic movement, **Gabriel Martínez Vera** sets out to discuss *wh*-extraction in Spanish clauses involving a *V+de+CP* sequence. With the help of the aforementioned theoretical assumptions, Martínez Vera derives the special properties exhibited by *V+de+CP* constructions with respect to extraction: the fact that subjects pattern with adjuncts and cannot be extracted from *de+CP* and the 'distance' effect (one extra level of embedding improves subject and adjunct extraction).

Guido Mensching & Franziska Werner focus on the extraction of *wh*-marked complements out of French subject DPs in direct interrogatives derived by complex inversion (Fr. *De quel livre connais-tu la fin?* "Of which book do you know the end?"), which show an increased degree of acceptability in comparison to similar interrogative structures based on the *est-ce que* strategy (Fr. *?*De quel linguiste est-ce que les parents ont déménagé à Chartres?* "(intended) Of which linguist did the parents move to Chartres?"). While on first sight these structures violate the subject-island constraint, the authors show that in complex inversion it is the subject itself which moves to the CP domain; prior to this step, there is internal reordering of constituents within the subject DP (with the complement of N reaching the edge of the subject DP), and the reordered DP subject moves as a whole. The subject-island constraint is thus only apparently, not effectively, violated.

Elena Soare discusses the grammar of a non-finite form specific to Romanian from a comparative Romance perspective, the supine, focusing on the status of the prepositional supine construction headed by the preposition *la* "at, to" in contexts in which it occurs as a complement to motion verbs. Soare shows that this construction encodes a goal of motion event and admits a continuation in which this event goal is not reached. By reviewing the properties of locative prepositions and of supine constructions, a unified account is proposed according to which the prepositional supine selects an eventive bare noun introducing a non-achieved meaning in the shape of a [-bounded] path.

Donca Steriade's contribution, also presented as a plenary talk, approaches the issue of cyclic inheritance without containment in Romanian perfects. Steriade

shows that the stress and segmental structure of Romanian perfects can be predicted from the perfect participle, which follows the accentual pattern of simple words; this requirement of stem identity is analysed as an instance of cyclic inheritance. The participle represents the base and the tensed perfect forms are generated as its derivative, without the base being contained, morphologically or syntactically, in its derivatives. The derivation of Romanian perfects is part of a bigger set of phenomena which show that Romanian morphology is a rich source of forms based on cyclic inheritance without containment.

Alina Tigău & Klaus von Heusinger put forward a novel analysis of Romanian ditransitives, a derivational account which builds on the internal make-up of the two internal arguments. Tigău & von Heusinger explore the relation between the clitic doubling of indirect objects and the differential object marking of direct objects, which interact in an interesting and unexpected way. It is shown that this interaction follows from the similar internal structure of DOM-ed direct objects and clitic doubled indirect objects, both of which carry a [Person] feature, hence competing for the same probe and incurring blocking effects. Direct object cliticization bypasses the intervention effects.

Andra Vasilescu takes issue with an extremely controversial problem in Romanian and comparative Romance linguistics: the status of the two passive forms of Romanian (the *be*-passive and the reflexive passive). By introducing a distinction between a syntactic passive and a pragmatic passive (= a structure which serves a different meaning but can be assigned a passive meaning via conventional implicature), Vasilescu argues that in old Romanian the reflexive passives functioned as syntactic passives (especially due to Old Church Slavonic influence), while *be*-structures served as temporal or copular constructions conventionally implicating the passive meaning. Under the overwhelming influence of the Western Romance languages after the 18th century, *be*-structures grammaticalized for the passive meaning and *se*-structures developed into impersonal presentatives.

Although theoretically and methodologically diverse, the papers reunited here manage to arrange in a coherent volume. The unity of the volume is given by its diversity; these chapters – individually and as a whole – show how the Romance languages contribute to a better understanding of issues which are relevant in the present-day linguistic landscape: language acquisition, *n*-words, ellipsis phenomena, focus and polarity, ditransitive constructions, grammaticalization theory, differential object marking, language ecology, extraction, event structure, (phonological) cyclicity, passives and many more. Similar to all the *Going Romance* and LSRL volumes of selected papers, this collection too is a testimony to the role played by the Romance languages in shaping our current understanding of linguistic theory.

Acknowledgements: Peer-review and organisation of the conference

We would like to express our utmost gratitude to the following colleagues who served as anonymous reviewers, some of them contributing more than one report: Gabriela Alboiu, Andrei A. Avram, Larisa Avram, Hugo Cardoso, Silvio Cruschina, Anamaria Fălăuș, Mihaela Gheorghe, Giuliana Giusti, Virginia Hill, Roland Hinterholz, Daniela Isac, Adam Ledgeway, Jonathan MacDonald, Ana Maria Martins, Gabriela Pană Dindelegan, Sandra Paoli, Anna Pineda Cirera, Ian Roberts, Johan Rooryck, Edward J. Rubin, Andrés Saab, Giampaolo Salvi, Oana Săvescu, Michelle Sheehan, Petra Sleeman, Ioana Stoicescu, Nigel Vincent and Hedde Zeijlstra.

We are also extremely indebted to Joe Salmons (series editor), Iman Sheydaei Baghdadeh, Anke de Looper, and Patricia Leplae for their support in the editing of this volume.

The 2017 edition was held in December 7–9 at the Faculty of Letters of the University of Bucharest, and organized by the University of Bucharest's Department of Linguistics in collaboration with the 'Iorgu Iordan – Al. Rosetti' Institute of Linguistics (Romanian Academy); the Research Institute of the University of Bucharest (ICUB) and the 'Alumni' Association of the University of Bucharest also contributed to the organization of this event.

Of the Going Romance steering committee, we are especially grateful to Roberta D'Alessandro and Frank Drijkoningen, who enthusiastically supported the organization of the conference in Bucharest.

The following linguists from Bucharest contributed to the scientific and practical organization of the event: Andrei A. Avram, Ana-Maria Barbu, Gabriela Bîlbîie, Raluca Brăescu, Alexandra Cornilescu, Blanca Croitor, Andreea Dinică, Mihaela Gheorghe, Ionuț Geană, Ion Giurgea, Emil Ionescu, Alexandra Loizo, Carmen Mîrzea Vasile, Isabela Nedelcu, Irina Nicula Paraschiv, Dana Niculescu, Gabriela Pană Dindelegan, Mihaela Tănase-Dogaru, Emanuela Timotin, †Marius Sala, Camelia Stan, Rodica Zafiu and Andra Vasilescu (as well as Adina Dragomirescu and Alexandru Nicolae). The following junior colleagues from the university and from the institute were extremely helpful during the event, and their contribution should be acknowledged: Adnana Boioc Apintei, Mădălina Botez Stănescu, Ștefania Costea, Valentina Cojocar, Roxana Dincă and Andreea Gruianu.

References cited

- Babyonyshev, Maria, Jennifer Ganger, David Pesetsky & Kenneth Wexler. 2001. The maturation of grammatical principles: Evidence from Russian unaccusatives. *Linguistic Inquiry* 32(1). 1–44. <https://doi.org/10.1162/002438901554577>
- Bošković, Željko. 2015. From the complex NP constraint to everything. *The Linguistic Review* 32. 603–669. <https://doi.org/10.1515/tlr-2015-0006>
- Bošković, Željko. 2016. Contextual phasehood and the ban on extraction from complements of lexical heads: When does X become a phase?. In Miyoko Yasui & Manabu Mizuguchi (eds.), *Phase theory and its consequences*, 5–39. Tokyo: Kaitakusha.
- D’Alessandro, Roberta & Tobias Sheer. 2015. Modular PIC. *Linguistic Inquiry* 46(4). 593–624. https://doi.org/10.1162/LING_a_00195
- Nunes, Jairo. 2014. Adjunct control and edge features. In Peter Kosta, Steven L. Franks, Teodora Radeva-Bork & Lilia Schürcks (eds.), *Minimalism and beyond: Radicalizing the interfaces*, 79–108. Amsterdam: John Benjamins. <https://doi.org/10.1075/lfab.11.04nun>
- Nunes, Jairo. 2016. Edge features on phase heads or moving elements?. *Florida Linguistics Papers* 3. 63–80.

The acquisition of verbal passives by Portuguese-speaking children

Some data from comprehension

Celina Agostinho and Anna Gavarró Algueró

Macau University of Science and Technology / Universitat Autònoma de Barcelona

Previous cross-linguistic research showed that verbal passives are delayed in child grammar. Moreover, Maratsos et al. (1985) found that actional passives elicit more adult-like results than non-actional passives in child English. Hirsch & Wexler (2006) proposed that the adult-like results children achieved with actional passives are due to a resultative adjectival passive analysis, unavailable for non-actional verbs. Alternatively, Snyder & Hyams (2015) proposed that this delay is due to the need of semantic coercion to passivize non-actional verbs. Here we present an experiment testing children's comprehension of short and long passives of actional and perception verbs in European Portuguese, a language with different auxiliaries for adjectival and verbal passives. The results replicate previous findings for English, despite the difference in auxiliary.

Keywords: verbal passive, adjectival passive, language acquisition, European Portuguese

1. Introduction

It has been observed in previous studies that the comprehension of the verbal passive is delayed in acquisition, in comparison with actives (see Maratsos et al. 1985; Hirsch & Wexler 2006 for English; Pierce 1992 for Spanish; Terzi & Wexler 2002 for Greek; Gavarró & Parramon 2017 for Catalan, among others). In addition, in a study on the English passive, Maratsos et al. (1985) found an asymmetry between actional and subject experiencer (or non-actional) verbs: children performed significantly better with passives of actional verbs (67% correct responses) than with passives of subject experiencer verbs (40% correct responses). This result has been replicated for English in subsequent studies (e.g., Hirsch & Wexler 2006; Orfitelli 2012).

Results such as these have been interpreted as evidence that the verbal passive is understood late, that is, that passive grammar is delayed until school age (Borer

& Wexler 1987; Wexler 2004). Borer & Wexler (1987) proposed that the development of passive grammar is subject to maturation. In their original proposal, Borer & Wexler (1987) attributed children's non-adult performance with the verbal passive to their inability to form A-chains (ACDH: A-chain Deficit Hypothesis). Later theoretical developments and empirical findings made this hypothesis untenable: Koopman & Sportiche (1991) argued that, even in simple active clauses, the subject is merged in the VP, from which it is A-moved to a higher position. The ACDH predicts that this movement is delayed in child grammar. However, Stromswold (1996), on the basis of evidence from negation and auxiliaries, showed that English-speaking children have no problems with subject raising.

This led to reformulations of the ACDH as linguistic theory changed. Namely, Wexler (2004) proposed the Universal Phase Requirement (UPR). Adopting the view that the verbal passive involves a non-phasal defective v (Chomsky 2001), Wexler (2004) claimed that children subject to UPR assume that v always defines a strong phase. Specifically, children have the adult v^* (a v head selecting an external argument and defining a phase), but they do not have v_{def} (a v head that does not select an external argument and does not define a phase). Instead, they have v^*_{def} (a v head that does not select an external argument but defines a phase). Around age 6, v_{def} becomes maturationally available and the child is able to derive verbal passives. This hypothesis predicts a delay not only with verbal passives, but also with raising-to-subject in *seem*-type contexts and unaccusatives.

An empirical challenge to the UPR comes from findings by Snyder et al. (1995) and Snyder & Hyams (2015). These studies showed that 2-year-old children perform at ceiling in the production of Italian and French 'formally (but not semantically) reflexive clitic constructions' (FRCCs), which are argued to be unaccusative and to resemble the verbal passive, as they involve an underlying object that raises to subject position, but do not have an external argument (see the response to Snyder et al. 1995 in Wexler 2004). Moreover, data from processing of Dutch intransitive verbs suggest that children differentiate between unaccusative and unergative verbs and that they have an underlying representation of unaccusatives that includes the internal argument (Koring et al. 2018). We set aside the issue of constructions other than the passive in the remainder of this chapter.

Wexler (2004) followed an independent claim by Borer & Wexler (1987) to explain the verb type asymmetry found in the comprehension of passives by English-speaking children: with a large subset of actional verbs (e.g., *comb*, *paint*, *tear*), but not with subject experiencer verbs (e.g., *see*, *hear*, *like*), English *be* + *participle* strings are potentially ambiguous between verbal passive and adjectival passive readings. This is because participles of actional verbs tend to make good adjectives, unlike participles of non-actional verbs, as illustrated by the contrast in (1)–(2):

- (1) Actional verbs
 - a. *the doll appears combed; the combed doll; combed though the doll was, Janie recombed her*
 - b. *the doll appears torn; the torn doll; torn though the doll was, John decided to keep her*
- (2) Non-actional verbs
 - c. **the doll appears seen; *the seen doll; *seen though the movie was, John decided to go again*
 - d. **the doll appears liked; *the liked doll; *liked though the doll was, John did not keep it*
(Borer & Wexler 1987: 135)

Under the assumption that children comprehend adjectival passives, but not verbal passives (which are ungrammatical under ACDH/UPR), they may assign an adjectival passive structure to verbal passives of actional verbs, but not to verbal passives of subject experiencer verbs, resulting in the observed asymmetry. A later study termed this “syntactic homophony” (abbreviated s-homophony) and defined it as follows: “A phrase α is an s-homophone of β if α and β have distinct structure but common pronunciation” (Babyonyshev et al. 2001: 7). This is illustrated in (3):

- (3) *The door was closed.*
 - i. Someone was closing the door.
 - ii. The door was in a closed state.

Borer & Wexler (1987) cited data from Horgan (1975) as indirect evidence for the early acquisition of adjectival passives. She analysed children’s production when describing pictures and observed that children’s short passives tend to be “after-the-fact observations” on the state of things. However, she noted that “since truncated passives were especially common in pictures with the agent deleted, these may be the result of the stimuli rather than the child’s strategy” (Horgan 1975: 94).

Moreover, the claim that subject experiencer verbs do not form adjectival participles in English was subject to criticism from early on. Namely, Weinberg (1987) argued that the formation of adjectival participles is not restricted to actional verbs, as shown by the examples in (4):

- (4) *a respected woman*
an admired man
an appreciated complement
a despised dictator
a loathed criminal
an expected retort
(Weinberg 1987: 178)

Subsequently, Hirsch & Wexler (2006), following Embick's (2004) tripartite typology of passives, proposed that 'pre-mature' children interpret verbal passives as resultative adjectival passives, whenever this is viable. They argued that this predicts the verb type contrast found in English: actional verbs typically have a target state, hence they form good resultative adjectival passives. Conversely, non-actional verbs typically do not involve a target state and do not form good resultative adjectival passives. As they pointed out, a prediction of this claim, pending confirmation, is that children will obtain lower comprehension scores with actional verbs that form poor resultative passives, and achieve better results with non-actional verbs that form good resultative passives.

An alternative hypothesis claims that children's delay with the verbal passive stems from the unavailability of the necessary mechanism to overcome a potential intervention effect in these constructions. Hyams & Snyder (2005) adopted Collins' (2005) Smuggling account of the passive and suggested that children over-generalize the Freezing Principle (Müller 1998) to structures that involve Smuggling (UFH: Universal Freezing Hypothesis). Smuggling becomes maturationally available around age 4, a change that affects only actional passives. Passives of subject experiencer verbs require further development. Snyder & Hyams (2015) adopted Gehrke & Grillo's (2007, 2009) account of passivization with stative verbs: semantic coercion is necessary, as these verbs lack a higher VP shell encoding the cause of a change of state, which is required for Smuggling in the passive. It is only around age 6 that semantic coercion becomes available. Hence, children have knowledge of verbal passive grammar by age 4, but this is obscured in the case of subject experiencer (stative) verbs due to the need of an additional mechanism to derive the passive.

However, the Smuggling account of the passive is controversial (e.g., Kiparsky 2013; Lima Júnior & Augusto 2015; Manzini 2017). In addition, Gehrke & Grillo's (2007, 2009) account of passivization with non-actional verbs relies heavily on semantic coercion, a process that is poorly understood (see also Snyder & Nguyen 2017 for an alternative view of semantic coercion in passivization).

Finally, the UFH is not incompatible with Borer & Wexler's (1987) suggestion that young English-speaking children may analyse verbal passives as adjectival passives. It is tenable that children up to age 4, subject to Universal Freezing, have adjectival interpretations of some verbal passives.

It is important to note that, although passive delay has been observed for English and other languages by an impressive number of studies, some studies have challenged the generalization that the verbal passive is delayed in child grammar. Namely, Crain et al. (2009) and O'Brien et al. (2006) attribute children's poor results with the verbal passive in previous studies to the experimental procedures used, which failed to pragmatically license the use of the passive. That is, according to this view, passive

delay reduces to an experimental artifact. Cross-linguistic research has also called into question the idea that the verbal passive is acquired late: early acquisition of the passive (around age 3–4) has been argued for in Sesotho (Demuth 1989), Italian (Volpato et al. 2016) and Brazilian Portuguese (Lima Júnior et al. 2018), among other languages. Crucially, both Sesotho and Brazilian Portuguese, like European Portuguese, display morphosyntactic distinctions between verbal and adjectival passives, and, while Italian short passives with the auxiliary *essere* may be either verbal or adjectival, passives with the auxiliary *venire* are unambiguously verbal.

In this context, the study of passives in European Portuguese is of special interest, as the s-homophony of verbal and adjectival passives does not hold. This chapter is organized as follows. In § 2, we discuss the relevant properties of EP and a previous study on the acquisition of the passive in this language. In § 3, we describe our experimental design and report our results. In § 4, we discuss these results in view of previous empirical studies and conclude the chapter.

2. The passive in European Portuguese

The verbal passive in EP is formed with the passive auxiliary *ser*, followed by the participial form of the main verb. This participle agrees in gender and number with the derived subject. As in English and in other Romance languages, the external argument may or may not be overtly expressed in a *by*-phrase (5).

- (5) *O exemplo foi corrigido* (por um falante nativo).
 the example AUX correct.PART.RG by a speaker native
 “The example was corrected (by a native speaker).”

(Duarte & Oliveira 2010: 401)

Adjectival passives, as illustrated in (6), may take the auxiliaries *ficar* and *estar*.¹

- (6) a. *O exemplo ficou corrigido*.
 the example AUX correct.PART.RG
 “The example got corrected.”

- b. *O exemplo está correcto*.
 the example AUX correct.PART.IRG
 “The example was correct.”

(Duarte & Oliveira 2010: 401)

1. EP shows double participle formation, that is, some verbs (e.g., *corrigir* ‘correct’, *limpar* ‘clean’, *entregar* ‘deliver’) have both a regular form of the participle (*corrigido*, *limpado*, *entregado*) and an irregular form (*correcto*, *limpo*, *entregue*) – RG and IRG, respectively, in the glosses. In some cases, the irregular form has been fully recategorized as an adjective. None of the verbs used in our experiment shows double participle formation.

Duarte & Oliveira (2010) extended Embick's (2004) tripartite typology of participles to EP to account for these data. Besides eventive participles, which enter the verbal passive construction, Embick (2004) distinguished between two subclasses of adjectival participles: resultative and stative. He argued that eventive participles have an agentive subcomponent that is absent from both subclasses of adjectival participles. Resultative participles share with eventive participles the presence of an eventive subcomponent. This gives rise to the requirement of a previous event that causes the result state denoted by the resultative participle. Conversely, stative participles denote pure states.

The data from EP, Duarte & Oliveira (2010) argued, support this typology. They suggested that Embick's (2004) typology is reflected in auxiliary selection: participles introduced by *ficar* are resultative, whereas participles introduced by *estar* are typically stative. As for eventive participles, they are introduced by the passive auxiliary *ser*.

Some diagnostics support these claims. The lack of an implicit agent in *ficar* and *estar* passives is shown by the ungrammaticality of agent-oriented adverbs in these constructions, as exemplified in (7). Conversely, verbal passives allow agent-oriented adverbs (8), as they have an agentive component.

- (7) a. **O exemplo ficou propositadamente corrigido.*
 the example AUX.RESULT purposefully correct.PART.RG
 "The example got purposefully corrected"
- b. **O exemplo está propositadamente correcto.*
 the example AUX.STATE purposefully correct.PART.IRG
 "The example was purposefully correct."
- (8) *O exemplo foi propositadamente corrigido.*
 the example AUX.EVENT purposefully correct.PART.RG
 "The example was purposefully corrected."

(following Duarte & Oliveira 2010: 401)

The lack of an eventive subcomponent in *estar* passives is shown by their ability to occur as complements of some change of state verbs. This is disallowed by resultative and eventive participles, as illustrated is (9):

- (9) a. *Construiu-se o bunker oculto.*
 built+CL the bunker hide.PART.IRG
- b. **Construiu-se o bunker ocultado.*
 built+CL the bunker hide.PART.RG
 "The bunker was built hidden." (Duarte & Oliveira 2010: 401)

In addition, as shown in (10), *estar* passives allow intensifiers, indicating that they involve a true adjective, unlike *ficar* and *ser* passives:

- (10) a. *O exemplo está correctíssimo.*
 the example AUX.STATE CORRECT.PART.IRG.SUP
 “The example was very correct.”
- b. **O exemplo ficou/foi corrigidíssimo.*
 the example AUX.RESULT/AUX.EVENT CORRECT.PART.RG.SUP
 “The example got/was very corrected.”

Hence, EP displays a morphosyntactic reflex of participle typology that is absent from English. Moreover, while Romance languages such as Spanish and Catalan also have the *ser/estar* distinction, they do not have an auxiliary specific to resultative adjectival passives.

To the best of our knowledge, there has been only one study on the acquisition of verbal passives in EP. Estrela (2013) replicated the passive delay and verb type asymmetry found in English: the results of her second experiment, which tested 3-, 4- and 5-year-old children on short and long passives of actional and non-actional verbs, show that children from the age of 4 perform at adult level with actional passives (79% correct responses, above the 75% cut off Estrela used to define adult level), whereas with non-actional passives 5-year-old children still have not reached adult level (64% correct responses). She also reported that children show similar performance on short and long passives of actional verbs. With non-actional verbs, children do better with short passives than with long passives, but this difference is not statistically significant (Estrela 2013: 185–186).

However, Estrela’s experiments present some flaws, which may have negatively impacted children’s performance. Some verbs were used inappropriately (e.g., in the sentence *O rapaz está a ser divertido* the main verb is used as a true transitive, with the meaning “amuse”, although *divertir* typically occurs with the reflexive pronoun *se* with the meaning “have fun”) or may not be known by younger children (e.g., *examinar* “examine”, *avistar* “see (at a distance)”); some of the actions are usually described using a periphrasis instead of a verb (*dar um abraço* instead of *abraçar* and *dar um beijo* instead of *beijar*), and the items were in the present simple, which is generally infelicitous in the EP verbal passive, due to the fact that the present simple does not have an ongoing interpretation in EP.

The main goal of the present study is to assess the acquisition path of the verbal passive by EP-speaking children, by testing children’s interpretation of verbal passives using a carefully designed method. We also consider some implications of our results for the hypotheses above, namely the UPR and the UFH.

3. Experimental task: Predictions, procedure and results

We used a two-choice sentence-picture matching task to assess passive acquisition in EP. The experiment tested short and long passives of actional verbs (*pentear* “comb”, *pintar* “paint”) and perception verbs (*ouvir* “hear”, *ver* “see”). Two control conditions testing actives with actional and perception verbs ensured that children understood the procedure and the scenarios. Children who scored fewer than 3 (out of 4) correct items in either of these two conditions were excluded from the study. Each condition was tested four times, with two items per main verb, for a total of 24 items.

Choice of verbs was largely determined by their imageability – hence the use of perception verbs in the conditions testing subject experiencer verbs. The actional verbs *pentear* “comb”, *pintar* “paint” and the perception verb *ver* “see” are also attested in the CEPLEXicon lexicon of Child European Portuguese (Santos et al. 2014), which was compiled from two corpora of child production and child-directed speech, the Santos corpus (Santos 2009; Santos et al. 2014) and the Freitas corpus (Freitas 1997).

It should be noted that in EP the verbs *ver* and *ouvir* can be used both as non-actional and actional verbs (“see”/“watch” and “hear”/“listen”). However, even when used as actional verbs, *ver* and *ouvir* typically do not include a result state in their event structure, and do not typically form good, uncoerced adjectival passives – see Examples (15)–(16) in § 4.

Under the UPR and assuming that a verbal passive must be strictly identical to an adjectival passive for s-homophony to hold, all verbal passives are predicted to be poorly comprehended by EP-speaking children until the age of 6–7, given the difference in auxiliary. Gavarró & Parramon (2017), however, argued that the conditions on s-homophony should be loosened, in view of data from Catalan suggesting that children acquiring this language resort to adjectival readings of verbal passives, despite the fact that verbal passives take the auxiliary *ser* and adjectival passives the auxiliary *estar*. Specifically, Gavarró & Parramon (2017) suggested that in a language such as Catalan (and by extension EP), in which the difference between verbal and adjectival passives resides only in auxiliary selection and is thus not prominent, children may still resort to adjectival interpretations of verbal passives (see also Oliva & Wexler 2018, who adopted this proposal for Spanish). If this is the case, EP-speaking children are predicted to pattern with English-speaking children, with significantly poorer performance on passives with *ver* “see” and *ouvir* “hear” than on passives with *pentear* “comb” and *pintar* “paint”. Moreover, children are predicted to show poorer performance on actional long passives than on actional short passives, given that *by*-phrases are highly restricted in adjectival passives (Borer & Wexler 1987; Hirsch & Wexler 2006; Oliva & Wexler 2018).

Under the UFH, passives with *ver* “see” and *ouvir* “hear” are predicted to be delayed until around age 6, as they require coercion in order to passivize, whether they are interpreted as actional or non-actional, given the absence of a result state. Passives of the actional verbs *pentear* “comb” and *pintar* “paint”, on the other hand, are predicted to elicit adult-like performance from the age of 4 years.

To be clear, the relevant contrast for testing the hypotheses described above is not actional verbs vs subject experiencer verbs. Rather, the contrast is between verbs that form good adjectival passives vs verbs that do not (under UPR), and between verbs that typically have a result state vs verbs that do not (under UFH).

The participants in our study were instructed to listen to very short scenarios about a family, listen to what a puppet (Benny) said at the end of each scenario (the test sentence), and choose the picture matching the test sentence. All passive items were semantically reversible. Given that EP presents participial agreement in gender and number, all the characters in each item had the same gender. Half of the items had female characters and the other half had male characters. In view of criticisms raised by O’Brien et al. (2006) to previous studies, all pictures included a third character to ensure the pragmatic felicitousness of a full *by*-phrase. The picture pairs were presented to participants side-by-side on an iPad with a 9.7” display. One of the pictures matched the adult interpretation of the test item, the other conveyed the theta-role reversal interpretation – see the example items in (11)–(12) and Figures 1–2. The order of presentation was pseudo-randomized, so that participants were not tested on the same condition or main verb twice in a row.

(11) *Item 11, Condition 3 – Actional short passives*

Experimenter: *Look, it’s the girl, the mother and the grandmother. They were getting ready to go out. Benny, what happened?*

Benny: *Então... A mãe foi penteada.*
then the mother was combed

Experimenter: *Which image shows what Benny said?*

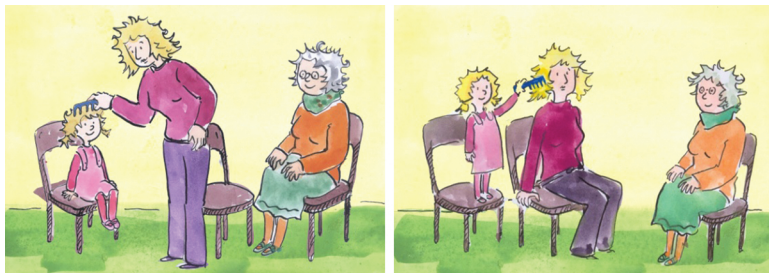


Figure 1. Picture pair for Item 11, *A mãe foi penteada* (“The mother was combed”)

(12) *Item 16, Condition 4 – Perception short passives*

Experimenter: *Look, it's the boy, the father and the grandfather. They were in the living room. One was reading, the other singing and the other listening. Benny, what happened?*

Benny: *Então... O menino foi ouvido.*
then the boy was heard

Experimenter: *Which image shows what Benny said?*



Figure 2. Picture pair for Item 16, *O menino foi ouvido* (“The boy was heard”)

(13) *Item 20, Condition 5 – Actional long passives*

Experimenter: *Look, it's the girl, the aunt and the grandmother. They were playing in the living room. Benny, what happened?*

Benny: *Então... A menina foi pintada pela tia.*
then the girl was painted by+the aunt

Experimenter: *Which image shows what Benny said?*



Figure 3. Picture pair for Item 20, *A menina foi pintada pela tia* (“The girl was painted by the aunt”)

(14) *Item 23, Condition 6 – Perception long passives*

Experimenter: *Look, it's the boy, the father and the grandfather. They were in the living room. The boy and the grandfather were playing blind man's buff. Look, there was cheating! Benny, what happened?*

Benny: *Então... O menino foi visto pelo avô.*
then the boy was seen by+the grandfather

Experimenter: *Which image shows what Benny said?*



Figure 4. Picture pair for Item 23, *O menino foi visto pelo avô* (“The boy was seen by the grandfather”)

All participants were monolingual speakers of EP. None of them had been diagnosed with a language delay or a language, hearing or cognitive impairment. Five children (four 3-year-olds and one 5-year-old) were excluded due to failure to answer correctly at least 3/4 items in each active sentence condition. In this contribution, we present results from 117 children between the ages of 3 and 6 years who were tested in several pre-schools in Lisbon, as well as a control group of 20 adults (see Table 1).

Table 1. Participants (137)

Group	N	Age range (mean)
3-year-olds	37 (13 girls, 24 boys)	3;0.08–3;11.24 (3;6)
4-year-olds	28 (20 girls, 8 boys)	4;0.00–4;11.04 (4;5)
5-year-olds	26 (13 girls, 13 boys)	5;0.11–5;11.14 (5;5)
6-year-olds	26 (13 girls, 13 boys)	6;0.10–6;11.26 (6;6)
Adults	20 (12 women, 8 men)	≥ 19

The results were coded for correctness and analysed using a Generalized Linear Mixed Model (GLMM). We report the results as follows: Table 2 and Figure 5 show the proportions of correct responses participants gave in each condition.

Table 2. Proportion of correct responses per condition

Condition	3 years	4 years	5 years	6 years	Adults
1. Actional actives	98.93%	99.29%	99.27%	100%	100%
2. Perception actives	97.3%	99.29%	99.27%	100%	98.8%
3. Actional short passives	95.06%	99.29%	97.77%	100%	100%
4. Perception short passives	42.05%	64.91%	72.52%	73.89%	98.8%
5. Actional long passives	63.79%	90.08%	92.18%	99.31%	100%
6. Perception long passives	33.43%	50.28%	47.88%	68.65%	96.2%

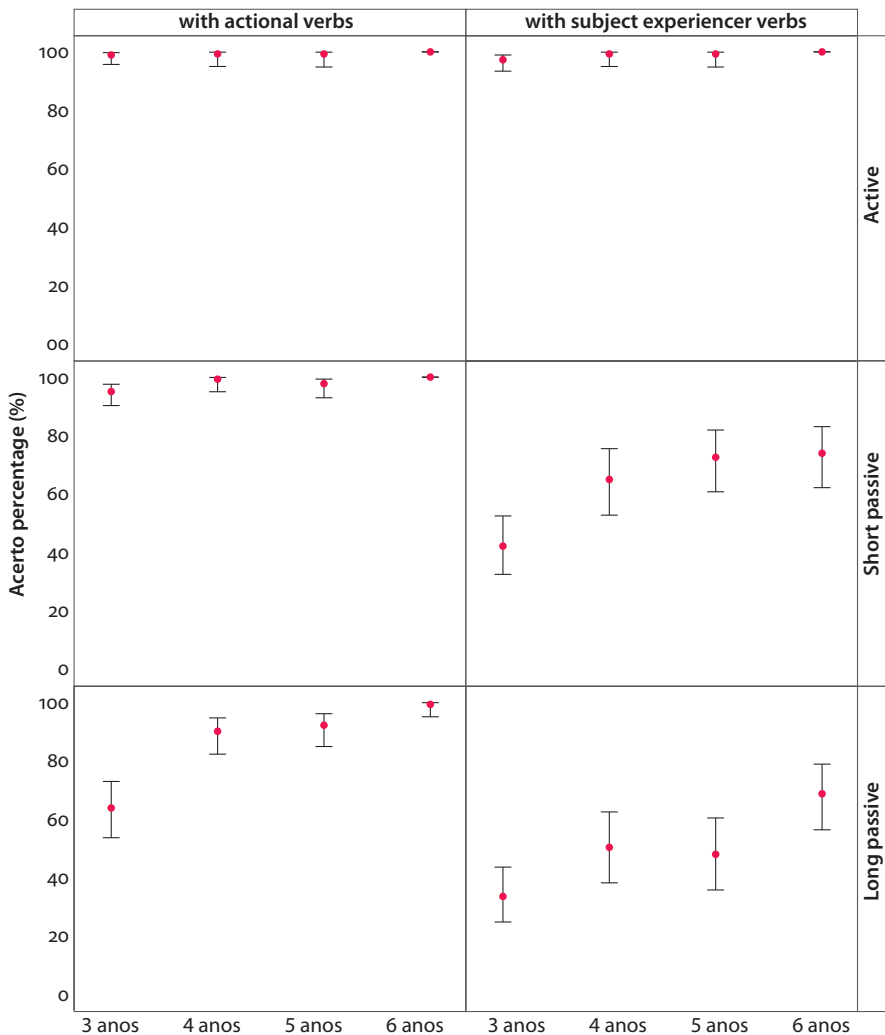


Figure 5. Proportion of correct answers per condition (with standard deviation)

The results show that, with actional short passives, all age groups performed at ceiling or nearly at ceiling: the 3-year-olds, who had the lowest score, showed a correctness rate of 95.06% (CI_{95%} = [90.25%, 97.56%]). There were no statistically significant differences between actional short passives and the corresponding actives. With perception short passives, 3-year-olds failed to perform above chance (42.05% CI_{95%} = [32.35%, 52.39%]). However, the results of the statistical analysis showed that all groups from age 4 performed above chance (at age 4, 64.91% CI_{95%} = [52.66%, 75.46%]; at age 5, 72.52% CI_{95%} = [60.67%, 81.86%]; at age 6, 73.89% CI_{95%} = [62.09%, 83.02%]). With actional long passives, all child groups performed above chance (at age 3, 63.79% CI_{95%} = [53.58%, 72.89%]; at age 4, 90.08% CI_{95%} = [82.24%, 94.69%]; at age 5, 92.18% CI_{95%} = [84.87%, 96.12%]; at age 6, 99.31% CI_{95%} = [95.07%, 99.91%]), whereas with perception long passives only 6-year-olds performed above chance (68.65% CI_{95%} = [56.3%, 78.82%]).

The results of the GLMM showed that there is a Sentence Type*Age Group interaction (F Value = 163.98; $p < 0.0001$). Collapsing the two verb types, all age groups performed above chance with short passives (at age 3, 69.32% CI_{95%} = [62.62%, 75.29%]; at age 4, 82.29% CI_{95%} = [75.63%, 87.43%]; at age 5, 84.85% CI_{95%} = [78.46%, 89.59%]; at age 6, 86.84% CI_{95%} = [80.80%, 91.19%]). With long passives of both verb types, 3-year-olds performed at chance level (48.62% CI_{95%} = [41.6%, 55.71%]), while *all* other child groups performed above chance (at age 4, 70.11% CI_{95%} = [62.12%, 77.04%]; at age 5, 70.32% CI_{95%} = [62.19%, 77.34%]; at age 6, 84.04% CI_{95%} = [77.5%, 88.96%]). There were statistically significant differences between actives and short passives (at age 3, t value = 7.52, $p < .0001$; at age 4, t value = 4.46, $p = 0.0006$; at age 5, t value = 4.13, $p = 0.0024$; at age 6, t value = 35.17, $p < .0001$) and between actives and long passives (at age 3, t value = 9.71, $p < .0001$; at age 4, t value = 5.43, $p < .0001$; at age 5, t value = 5.36, $p < .0001$; at age 6, t value = 33.24, $p < .0001$).

These data suggest that, when actional and perception verbs are collapsed and both short and long passives are considered, 3- to 6-year-old children comprehend actives better than passives. The results also indicate that EP-speaking children fully comprehend short actional passives, as they perform at ceiling with these sentences, and there were no statistically significant differences with regard to the corresponding actives. With short perception passives, 3-year-olds performed at chance, whereas older children performed above chance, but with statistically significant differences with respect to the corresponding actives. Hence, EP seems to display a contrast between actional and perception verbs, similar to English. As for long passives, all child groups performed above chance with actional verbs, whereas with perception verbs 4- and 5-year-olds performed at chance level and 3-year-olds at below-chance level.

We also observed a lexical contrast in the case of passives of perception verbs. Table 3 and Figure 6 show the proportions of correct answers children gave with each perception verb on short and long passives.

Table 3. Proportion of correct answers per perception verb

Sentence type	Main verb	3 years	4 years	5 years	6 years
Short passive	<i>ouvir</i>	34.20%	50.25%	53.01%	49.97%
	<i>ver</i>	48.20%	81.20%	92.70%	95.92%
Long passive	<i>ouvir</i>	32.70%	31.47%	15.57%	44.71%
	<i>ver</i>	31.31%	68.85%	80.69%	91.90%

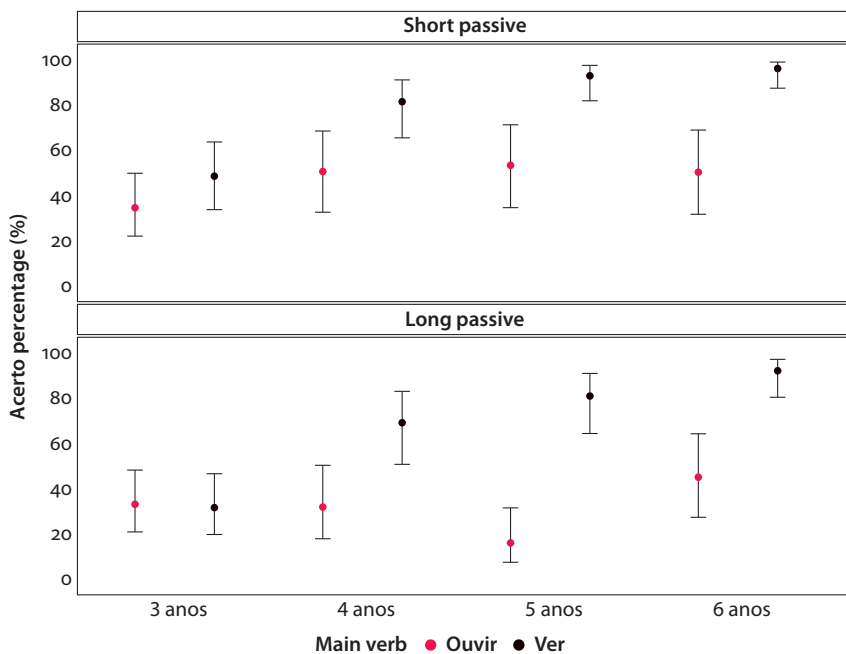


Figure 6. Proportion of correct answers per perception verb (with standard deviation)

A GLMM was set up and applied only to the data from short and long passives of perception verbs. The results of this analysis show that there is a Length*Age Group*Main Verb interaction (F value = 4.55, $p < 0.0001$). However, only the 5- and 6-year-olds showed statistically significant differences between the two main verbs. The 5-year-olds showed statistically significant differences for short passives (t value = -4.20 , $p = 0.0035$) and long passives (t value = -5.62 , $p < 0.0001$) of *ouvir* and *ver*. Similarly, the 6-year-old group showed statistically significant differences

for short passives (t value = -4.69 , $p = 0.0004$) and long passives (t value = -4.55 , $p = 0.0008$) of *ouvir* and *ver*. For actional passives, no lexical contrast was observed in any age group.

4. Discussion and conclusion

The results of our experiment show that active sentences are better understood than passive sentences for all age groups, if we collapse actional and perception verbs and short and long passives. However, EP-speaking children as young as 3 performed at ceiling with short actional passives, and there were no statistically significant differences between actives and short passives of actional verbs in any age group. With long actional passives, 3-year-olds performed above chance, and 4-year-olds performed nearly at ceiling. Hence, although EP has different auxiliaries for verbal (*ser*) and adjectival (*ficar, estar*) passives, children from the age of 3 are seemingly able to comprehend verbal passives with actional verbs.

With perception verbs, for short passives children from the age of 4 performed above chance, while children at the age of 3 performed at chance. With long passives, 3-year-olds performed below chance, 4- and 5-year-olds at chance and 6-year-olds above chance. These results indicate that EP-speaking children display a contrast between actional and perception verbs.

Finally, passives of perception verbs showed a contrast between the verbs *ouvir* and *ver*: 5- and 6-year-olds performed better with *ver* than with *ouvir*, for both short and long passives. At ages 3 and 4, however, there were no statistically significant differences between the two verbs. A similar contrast was observed in some studies on passive acquisition in English (Maratsos et al. 1985; Fox & Grodzinsky 1998; O'Brien et al. 2006). Fox & Grodzinsky (1998) found that children between the ages of 3;6 and 5;5 performed at chance with long passives of *see* (55% correct responses) and below chance with long passives of *hear* (25% correct responses), and attributed the difference to the availability of an actional interpretation of *see*. Maratsos et al. (1985) observed that, of all the mental verbs they tested (*see, hear, remember, like, love, hate*), *hear* was the most difficult for children at the ages of 4, 5 and 7. Children's performance with mental verbs was also more heterogeneous than with actional verbs (*wash, kiss, push, kick, find, hold*). Indeed, children's correctness rates showed no significant differences across actional verbs. O'Brien et al. (2006) also found that 4-year-olds performed poorly on short passives of *hear*. The same subjects reportedly performed significantly better on short passives of *see*. Hence, the authors excluded all items with *hear* from the analysis of the results and, like Maratsos et al. (1985), they do not present children's correctness rates with this

verb. Snyder & Hyams (2015) argued against this decision, on the grounds that the items with *see* (the only remaining typically non-actional verb) used this verb in an actional sense ('discover' or 'find').

In EP, the verb *ver* "see" may in some contexts receive an actional reading ('discover' or 'find'). In particular, the items we used to test this verb may have led young children to interpret *ver* as an actional verb ('discover') – see the example item in (14). By hypothesis, under the UFH the availability of this reading could lead to better results, given that, it is assumed, semantic coercion is rendered unnecessary. Nonetheless, young children's performance on passives of *ver* remains poorer than their performance on passives of *pentear* "comb" and *pintar* "paint", which casts doubt on this interpretation of the results.

In addition, both *ver* "see" and *ouvir* "hear" may occur in adjectival passives. The examples in (15)–(16) illustrate this point. The (abbreviated) example in (16) is taken from the corpus CETEMPúblico. Notice that, in these examples, the verbs *ver* and *ouvir* have agentive readings:

- (15) *Está/ficou vista a exposição!*
 AUX.STATE/AUX.RESULT see.PART the exhibition
 "The exhibition has been seen!"
- (16) ... *ouvido um disco, estão ouvidos os próximos dez.*
 hear.PART one album AUX.STATE hear.PART the next ten
 "... having listened to one album, you've listened to the next ten."

Both (15) and (16) are somewhat coerced. Assuming the UPR, children may interpret verbal passives with *ver* as adjectival passives. If young children ignore the auxiliary *ser* or do not yet distinguish it from the adjectival passive auxiliaries (*estar/ficar*), they may assign an adjectival interpretation to passives of *ver* (see Gavarró & Parramon 2017; Oliva & Wexler 2018). However, these adjectival readings are also available with *ouvir*, as shown in (16), hence the fact that children perform more poorly on passives of *ouvir* than on passives of *ver* remains unexplained.

The results from this experiment allow us to conclude that EP-speaking children have difficulties with passives of (at least some) subject experiencer verbs until school age. Only 6-year-olds performed above chance with long passives of perception verbs. This contrasts with their performance on long passives of actional verbs: as mentioned above, 3-year-olds performed at above-chance level with long actional passives.

Adult-like performance with passives of actional verbs is expected under UFH only from age 4, provided the actional verb has a result state, as is the case in our experiment. Lack of apparent passive delay with short passives of actional verbs is expected if children are interpreting verbal passives as adjectival and ignoring the

difference in auxiliary, as proposed by Gavarró & Parramon (2017) for Catalan. However, above-chance performance on actional long passives at early stages (age 3) is unexpected, as the *by*-phrase signals that the sentence is a verbal passive (Borer & Wexler 1987; Hirsch & Wexler 2006; Oliva & Wexler 2018). Experimental work directly addressing the interpretation that EP-speaking children assign to verbal and adjectival passives could shed further light on this issue.

Acknowledgments

The authors wish to thank the participants in the reported experiment and the schools where testing took place. We are also grateful to the audience of Going Romance 31 in Bucharest and to the reviewers of WoSSP 15 in Barcelona for comments on earlier versions of this work; to Anna Espinal and Ester Boixadera (SEA-UAB) for their help with the statistics; and to Ana Lúcia Santos (Universidade de Lisboa) for discussion and support in data gathering. Justyna Kamykowska produced the drawings used in the experiment.

Funding

This research was supported by projects FFI2014-56968-C4-1-P and FFI2017-87699-P of the Universitat Autònoma de Barcelona.

References

- Babyonyshev, Maria, Jennifer Ganger, David Pesetsky & Kenneth Wexler. 2001. The maturation of grammatical principles: Evidence from Russian unaccusatives. *Linguistic Inquiry* 32(1). 1–44. <https://doi.org/10.1162/002438901554577>
- Borer, Hagit & Kenneth Wexler. 1987. The maturation of syntax. In Tom Roeper & Edwin Williams (eds.), *Parameter setting*, 123–172. Dordrecht: Reidel. https://doi.org/10.1007/978-94-009-3727-7_6
- CETEMPúblico (Corpus de Extractos de Textos Electrónicos MCT/Público). <http://www.linguateca.pt/CETEMPUBLICO/>
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
- Collins, Chris. 2005. A smuggling approach to the passive in English. *Syntax* 8(2). 81–120. <https://doi.org/10.1111/j.1467-9612.2005.00076.x>
- Crain, Stephen, Rosalind Thornton & Keiko Musarugi. 2009. Capturing the evasive passive. *Language Acquisition* 16(2). 123–133. <https://doi.org/10.1080/10489220902769234>
- Demuth, Katherine. 1989. Maturation and the acquisition of the Sesotho passive. *Language* 65(1). 56–80. <https://doi.org/10.2307/414842>

- Duarte, Inês & Fátima Oliveira. 2010. Participípios resultativos. In Ana Maria Brito, Fátima Silva & João Veloso e Alexandra Fiéis (eds.), *Textos Seleccionados do XXV Encontro da Associação Portuguesa de Linguística*, 397–408. Porto: APL.
- Embick, David. 2004. On the structure of resultative participles in English. *Linguistic Inquiry* 35(3). 355–392. <https://doi.org/10.1162/0024389041402634>
- Estrela, Antónia. 2013. A aquisição da estrutura passiva em Português Europeu. Doctoral dissertation, Universidade Nova de Lisboa. https://doi.org/10.14195/978-989-26-0714-6_2
- Fox, Danny & Yosef Grodzinsky. 1998. Children's passive: A view from the *by*-phrase. *Linguistic Inquiry* 29(2). 311–332. <https://doi.org/10.1162/002438998553761>
- Freitas, Maria João. 1997. Aquisição da Estrutura Silábica do Português Europeu. MA thesis, University of Lisbon.
- Gavarró, Anna & Xavier Parramon. 2017. On the adjectival interpretation of passives in acquisition. *Lingua* 185. 11–24. <https://doi.org/10.1016/j.lingua.2016.07.003>
- Gehrke, Berit & Nino Grillo. 2007. Aspects on passives. In Sylvia Blaho, Eric Schoorlemmer & Luis Vicente (eds.), *Proceedings of ConSOLE XIV*, 121–142. Leiden: SOLE.
- Gehrke, Berit & Nino Grillo. 2009. How to become passive. In Kleantes K. Grohmann (ed.), *Explorations of phase theory: Features, arguments, and interpretation at the interfaces*, 231–268. Berlin: De Gruyter Mouton. <https://doi.org/10.1515/9783110213966.231>
- Hirsch, Chris & Kenneth Wexler. 2006. Children's passives and their resulting interpretation. *The Proceedings of the inaugural conference on Generative Approaches to Language Acquisition – North America*, 125–136. University of Connecticut Occasional Papers in Linguistics 4.
- Horgan, Dianne. 1975. Language development: A cross-methodological study. Doctoral dissertation, University of Michigan.
- Hyams, Nina & William Snyder. 2005. Young children never smuggle: Reflexive clitics and the universal freezing hypothesis. Paper presented at the 30th annual Boston University Conference on Language Development (BUCLD 30), Boston, MA, November 4–6.
- Kiparsky, Paul. 2013. Towards a null theory of the passive. *Lingua* 125. 7–33. <https://doi.org/10.1016/j.lingua.2012.09.003>
- Koopman, Hilda & Dominique Sportiche. 1991. the position of subjects. *Lingua* 85. 211–259. [https://doi.org/10.1016/0024-3841\(91\)90022-W](https://doi.org/10.1016/0024-3841(91)90022-W)
- Koring, Loes, Pim Mak, Iris Mulders & Eric Reuland. 2018. Processing intransitive verbs: How do children differ from adults?. *Language Learning and Development* 14(1). 72–94. <https://doi.org/10.1080/15475441.2017.1380528>
- Lima Júnior, João & Marina Augusto. 2015. Is 'smuggling' really necessary? The most recent analyses of passive sentences reconsidered in terms of phasehood and cyclic movement. *ReVEL* 9. 62–91.
- Lima Júnior, João, Letícia Sicuro Corrêa & Marina Augusto. 2018. Favorable processing conditions in the production of passive sentences by Brazilian Portuguese-Speaking children. In Anna Gavarró (ed.), *On the acquisition of the syntax of Romance*, 91–115. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.62.05lim>
- Manzini, Rita. 2017. Passive, smuggling and the *by*-phrase. In Nicholas LaCara, Keir Moulton & Anne-Michelle Tessier (eds.), *A schrifft to fest Kyle Johnson*, 233–244. Linguistics Open Access Publications 1. Amherst, MA: University of Massachusetts Amherst.
- Maratsos, Michael, Dana Fox, Judith Becker & Mary Anne Chalkley. 1985. Semantic restrictions in children's passives. *Cognition* 19(2). 167–191. [https://doi.org/10.1016/0010-0277\(85\)90017-4](https://doi.org/10.1016/0010-0277(85)90017-4)
- Müller, Gereon. 1998. *Incomplete category fronting*. Dordrecht: Kluwer. <https://doi.org/10.1007/978-94-017-1864-6>

- O'Brien, Karen, Elaine Grolla & Diane Lillo-Martin. 2006. Long passives are understood by young children. In David Bamman, Tatiana Magnitskaia & Colleen Zaller (eds.), *Proceedings of the 30th annual Boston University Conference on Language Development (BUCLD 30)*, 441–451. Somerville, MA: Cascadilla Press.
- Oliva, Jesús & Kenneth Wexler. 2018. The acquisition of Spanish passives: The comparison between subject experiencer versus actional verbs and direct aspectual semantic evidence for the adjectival interpretation. In Anna Gavarró (ed.), *On the acquisition of the syntax of Romance*, 57–89. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.62.04oli>
- Orfitelli, Robyn. 2012. Argument intervention in the acquisition of A-movement. Doctoral dissertation, UCLA.
- Pierce, Amy. 1992. The acquisition of passives in Spanish and the question of A-chain maturation. *Language Acquisition* 1. 55–81. https://doi.org/10.1207/s15327817lao201_3
- Santos, Ana Lúcia. 2009. *Minimal answers: Ellipsis, syntax and discourse in the acquisition of European Portuguese*. Amsterdam: John Benjamins. <https://doi.org/10.1075/lald.48>
- Santos, Ana Lúcia, Aida Cardoso & Maria João Freitas. 2014. CEPLEXicon – A lexicon of child European Portuguese. ELRA ID: ELRA-L0094.
- Santos, Ana Lúcia, Michel Génèreux, Aida Cardoso, Celina Agostinho & Silvana Abalada. 2014. A corpus of European Portuguese child and child-directed speech. In Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk & Stelios Piperidis (eds.), *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC 2014)*, 1488–1491. Reykjavik: European Language Resources Association.
- Snyder, William & Emma Nguyen. 2017. It's hard to coerce: A unified account of raising-past-experiencers and passives in child English. Paper presented at the 13th Generative Approaches to Language Acquisition conference (GALA 13), Palma de Mallorca, Spain, September 7–9.
- Snyder, William & Nina Hyams. 2015. Minimality effects in children's passives. In Elisa Di Domenico, Cornelia Hamann & Simona Matteini (eds.), *Structures, strategies and beyond: Essays in honour of Adriana Belletti*, 343–368. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.223.16sny>
- Snyder, William, Nina Hyams & Paola Crisma. 1995. Romance auxiliary selection with reflexive clitics: Evidence for early knowledge of unaccusativity. In Eve V. Clark (ed.) *Proceedings of the 26th Annual Child Language Research Forum*, 127–136. Stanford, CA: CSLI Publications.
- Stromswold, Karin. 1996. Does the VP-internal stage really exist?. Paper presented at the 21st Annual Boston University Conference on Language Development (BUCLD 26), Boston, MA, November 2–4.
- Terzi, Arhonto & Kenneth Wexler. 2002. A-chains and s-homophones in children's grammar: Evidence from Greek passives. *North East Linguistics Society* 32(2). 519–538.
- Volpato, Francesca, Laura Verin & Anna Cardinaletti. 2016. The comprehension and production of verbal passives by Italian preschool-age children. *Applied Psycholinguistics* 36. 901–931. <https://doi.org/10.1017/S0142716415000302>
- Wexler, Kenneth. 2004. Theory of phasal development: Perfection in child grammar. *MITWPL* 48. 159–209.
- Weinberg, Amy. 1987. Comments on Borer and Wexler. In Thomas Roeper & Edwin Williams (eds.), *Parameter setting*, 173–187. Dordrecht: Reidel Publishing. https://doi.org/10.1007/978-94-009-3727-7_7

Plus in the French negative system

A presuppositional and non-quantificational n-word

Pascal Amsili and Claire Beyssade

Université Paris Diderot / Université Paris 8

This contribution focuses on the place of *plus* (“no more”) in the system of negation in French. Is it merely the presuppositional counterpart of *pas* (“not”), or does it belong to the class of n-words, such as *personne* (“nobody”) and *rien* (“nothing”), which can combine with each other and give rise to negative concord readings? We show that when *plus* co-occurs with an n-word, it brings different semantic contributions depending on the place it occupies. This leads us to assume that two instances of *plus* have to be distinguished: one that incorporates a sentential negation and the other which has only a presuppositional contribution. These two instances differ in assertive content, but convey the same presuppositional content. The chapter shows how to derive the observed interpretations in each kind of configuration.

Keywords: presupposition, quantification, n-words, negative concord, phrasal negation, sentential negation

1. Introduction

It is generally believed that *plus* belongs to the system of negation in French. When used with *ne* it forms a negative sentence, and in this sense it is similar to *pas* and the French n-words (1).¹ It can also combine in a sentence with other negative words like *rien* (n-thing) or *personne* (n-body) and give rise to a negative concord (NC) reading.

From this viewpoint *plus* is comparable with n-words and different from *pas*, since *pas* does not combine easily with an n-word and, when it does, gives rise only to a double negation (DN) reading. However, *plus* displays several properties which

1. Convention: n-words like *personne* can be glossed “anybody” or “nobody” depending on the interpretation. To avoid confusion in the glosses, we have adopted the approach of Zeijlstra (2009) and translate n-words with a specific notation: ‘n-body’ for *personne*.

distinguish it from the other n-words: it is presuppositional, not quantificational, it cannot be denied by *pas* (2), and it can occupy various positions in the sentence, either just after the verb, or before an n-word (3).

- (1) *Jean ne dort pas / plus.*
 Jean NE sleep not n-more
 “Jean doesn’t sleep (anymore).”
- (2) a. **Jean ne voit pas plus.*
 Jean NE sees not n-more
 “(intended) Jean doesn’t see anymore.”
 b. *Jean ne voit pas personne.*
 Jean NE sees not n-body
 “Jean doesn’t see anybody.”
- (3) *Plus personne n’ est venu.*
 n-more n-body NE is come
 “Nobody else [*lit.* not anymore person] has come.”

So it seems that the place of *plus* within the system of negation in French is not immediately obvious: should it be seen as contributing a sentential negation like *pas* or rather providing the same kind of negation as (canonical) n-words?

To answer this question, we briefly present the system of negation in French (§ 2). Then we recall what characterizes n-words (§ 3). To decide how well *plus* fits into this class, we provide a description of the uses of *plus* in French. The description suggests that, when *plus* co-occurs with an n-word in a sentence, its semantic contribution depends on the place it occupies (§ 4). At the semantic level, the relative scope of *plus* with respect to the n-word changes the presupposition of the utterance. We then show that making a distinction between two instances of *plus*, one which incorporates a sentential negation and one which has only a presuppositional contribution, allows us to derive the observed interpretations (§ 5).

2. Negation in French

The lexicon of negation is composed of the particle *ne*, the negation *pas*, and various n-words. One well known property of negation in French is that it forms sentential negation by means of two negative markers (it is called ‘Embracing Negation’ by Zeijlstra (2009)): the preverbal particle *ne* and the negation adverb *pas* (4a). A second interesting aspect is that when an n-word is used, canonical negation is formed without the negative adverb *pas*: the preverbal particle *ne* combines with the n-word to obtain a negative interpretation (4b).

- (4) a. *Pierre ne mange pas.*
 Pierre NE eats not
 “Pierre does not eat.”
- b. *Pierre ne mange rien.*
 Pierre NE eats n-thing
 “Pierre does not eat anything.”

These examples are sometimes considered in the literature as instances of negative concord (NC), since two elements together contribute only one negation but in our view this is rather misleading since the negative particle *ne* has entirely lost its negativity in contemporary French (it only survives in idioms like *je ne sais* (“I don’t know”)); we follow Zeijlstra (2009) in considering that *ne* does not participate in concord.

When more than one negative element combines with *ne*, we can distinguish two cases: either an arbitrary number of n-words combine with *ne* but without the negation adverb *pas*, or both *pas* and (at least) one n-word co-occur in the same sentence. We will now consider the various interpretations of these different cases.

Several n-words but no *pas*. When several n-words have to appear in the same proposition, only one instance of *ne* is necessary, and *pas* is not normally used:

- (5) a. *Personne n’ a rien dit.*
 n-body NE has n-thing said
- b. *Personne n’ a plus jamais rien dit à personne.*
 n-body NE has n-more n-ever n-thing said to n-body

The most frequent interpretation for such examples is a typical negative concord: the sentence is interpreted as containing only one logical negation (6a).² In addition to this preferred reading, a multiple negation reading can be obtained where each n-word contributes one negation (6b).

- (6) Interpretations for (5a) (P for *personne*, C for *choses* “things” and D for *dire* “say”):
- a. NC: $\neg\exists x (Px \wedge \exists y (Cy \wedge Dxy))$
- b. DN: $\neg\exists x (Px \wedge \neg\exists y (Cy \wedge Dxy))$

2. A reasonably good paraphrase in NC can be obtained by translating one n-word with a negation (*personne* → *nobody*) and all the other n-words by their negative polarity version (*personne* → *anybody*):

- a. *Personne n’a rien dit.* = (5a), NC
 “Nobody said anything”
- b. *Personne n’a plus jamais rien dit à personne.* = (5b), NC
 “Nobody ever said anything to anybody anymore”

This double negation reading is much rarer, but can be forced by intonation, or by specific syntactic or semantic contexts (Corblin et al. 2004). For instance, as an answer to a negative question, (5a) and the corresponding fragment answer in (7) can easily be interpreted as a DN. The dialogue situation in (8) also clearly favours a DN reading.

- (7) A: *Qui n'a rien dit ?*
 "Who said nothing?"
 B: *Personne.*
 "Nobody."
- (8) A: *Jean n'a rien dit.*
 "Jean said nothing."
 B: *C'est faux! Personne n'a rien dit !*
 "It is not true! No one said nothing!"

When more than two n-words appear together, as in (5b), it is clearly still possible to get a concord reading (9a), and also a DN reading (9b), in which two n-words become negative, while the others remain in a concord relation; in principle other multiple negation (MN) readings could be obtained, such as (9c), but the cognitive load required to process three negations within one single proposition is probably sufficient to explain why such examples are not found in real life situations.

- (9) Interpretations for (5b)
- | | | |
|----|---|----|
| a. | Nobody ever said anything to anybody anymore. | NC |
| b. | Nobody ever said anything to nobody anymore. | DN |
| c. | Nobody ever said nothing to nobody anymore. | MN |

Sentential negation *pas* along with one additional n-word. All authors agree that such combinations are very rare, but quite a number of exceptions have also been documented, as illustrated in (10).

- (10) a. *Ce n' est pas rien.*
 It_{NE} is not n-thing
- b. *Il ne va pas nulle part, il va à son travail.*
 He_{NE} goes not n-where, he goes to his work
 "He is not going nowhere, he is going to work." (Muller 1991)

The most frequent interpretation of such examples is that of double negation (DN): for instance, (10a) is interpreted as meaning *it is not the case that it is nothing*, and as pointed out by Muller (1991) and Corblin et al. (2004), it also has a special meaning effect, introducing the value of 'something important'.

An additional interpretation seems to be available in Standard French, very similar to what is called negative concord in most Romance languages: a mononegative

interpretation that we will call, following Larrivée (2016), Negative Doubling (ND). We do not use the term Negative Concord because it would be confusing, since all the literature insists on the fact that *pas* in French does not participate in NC (De Swart & Sag 2002: 403, among others). In a corpus study, we found a number of interesting cases (11). However the status of such very rare examples is still under discussion, from dysfluencies to archaisms, from regionalisms to pragmatic marking; we refer the reader to the recent study by Larrivée (2016) on this question which we will leave aside in the rest of this chapter.

- (11) a. *Marine Le Pen ment aux Français, elle n' a pas aucune*
 Marine Le Pen lies to French she NE has not no
réponse concrète.
 answer concrete
 “Marine Le Pen is lying to the French, she hasn’t got any concrete answer.”
France Inter, 10.3.2011, 7h52, in (Larrivée 2016)
- b. – *mais je sais pas rien faire mon oncle ...*
 – but I know not n-thing do my uncle ...
 “but I don’t know how to do anything (= I am unable to do anything),
 uncle ...”
Céline L-F, Mort à crédit, 1936

At this point, we can summarize the observations about the combinations of negative terms in French. Embracing negation requires that sentential negation is marked by both the particle *ne*³ (quite often but not always dropped in informal registers) and either *pas* or an n-word. When several negative terms appear in the same clause, we distinguish two cases. (A) When two or more n-words co-occur, at most one occurrence of *ne* is required to obtain a well formed sentence, and the sentence is preferably interpreted as NC, but can also be, context permitting, interpreted as DN.⁴ (B) The combination of *pas* with one or more n-words is generally considered as ill-formed, but can still be found in rare cases. Generally speaking, the interpretation is a double negative, but a small number of cases interpreted mononegatively have been found (ND). Some authors include *plus* in the list of n-words (termed *semi-négations* (“half-negations”) by Muller (1991)), while others do not (Corblin & Tovenà 2003). We address this issue in the following section.

3. It is well known that *ne* can be and is indeed quite often dropped in many registers of contemporary French. A large majority of the examples that are presented in this chapter are indeed quite acceptable without *ne*. However, since we focus here on Standard French, where the absence of *ne* has an impact solely on register, we prefer to insert it consistently.

4. Not exactly all n-words can combine with one another: the combinatorial possibilities are thoroughly described by Muller (1991).

3. The class of n-words

We present in this section various definitions that have been proposed for the class of n-words in general or half-negations in French, and we try to determine how well the distributional properties of *plus* make it a typical or marginal member of those classes.

3.1 “Half negations” (Muller 1991)

In his work on negation in French, Muller (1991: 53 et seq.) proposes an inventory of the vocabulary of negation and distinguishes on the one hand syntactic negation operators that are used to build, from a positive statement, the corresponding negative statement (as does *ne ... pas*) and on the other hand negative words, like *incroyable* “unbelievable” or *rarement* “rarely” which can enter a natural paraphrase relation with a negative statement. For any negative word *X*, there is a term *Y* such that *X* gets as a natural paraphrase *ne...pas Y* (Muller 1991: 56). This paraphrase relation can have a morphologic support: *in-croyable* = *pas croyable* “not believable” or not: *rarement* = *pas souvent* “not often”.

Among those negative words, Muller (1991: 250 et seq.) distinguishes a subset, which he calls half-negations and which he analyses as amalgamations of a negation and an operator *Y*. The property that distinguishes a half-negation from another negative word is that a half-negation *X* (e.g., *rien* “nothing”) combines with *ne* and without *pas* exactly in the same contexts as those where its positive counterpart *Y* (e.g., *quoi que ce soit/quelque chose* “anything”) combines with *pas* (12).

- (12) *Jean ne voit rien* (X) = *Jean ne voit pas quoi que ce soit* (Y)
 Jean NE sees nothing (X) Jean NE sees not anything (Y)

According to Muller (1991: 250, 281 et seq.), the class of half-negations is made up of indefinite pronouns (*personne, rien*), determiners (*aucun, nul*) which can also have pronominal uses, adverbs (*jamais, nulle part, nullement, aucunement, plus, guère*), the conjunction *ni* and the restrictive adverb *que*. This class is very heterogeneous, grouping terms of various syntactic categories. According to Muller (1991), the members of the class share several other properties:

They can be used alone, in elliptical contexts for instance, or as an answer, and receive a negative interpretation (13) (which is not the case for *pas*).

- (13) A. *Est-ce que Jean est absent ?*
 “Is Jean absent?”
 B. *Jamais / # Pas.*
 “Never / Not.”

They combine with *ne* and without *pas*, except in (rare) cases of double negation reading (14).

- (14) a. *Jean n'aime personne.*
 "Jean loves no one."
 b. *Jean n'aime pas personne. Il s'aime lui-même.*
 "Jean loves not no one. He loves himself."

They can be used with a positive interpretation: *aucun* "no" in (15) means *un* "one/any".

- (15) *Au malaise que lui apporte le paysan, il refuse de donner aucun sens religieux.*
 "To the embarrassment that the peasant brings him, he refuses to give any religious meaning." J. Favret-Saada (*Les Mots, la Mort, les Sorts*, p. 18)

They can be combined with one another, giving rise to a mononegative interpretation, called concord reading (16a) = (16b).

- (16) a. *Personne ne dit jamais rien.*
 n-body NE says n-ever n-thing
 b. *Personne ne dit à quelque moment que ce soit quoi que ce soit.*
 nobody NE says at any moment anything
 "Nobody ever said anything."

To sum up, the crucial thing is that in certain configurations n-words can induce a semantic negation, but in other configurations they cannot do so. The definition proposed by Muller (1991) is meant only for French negative words, but the class of n-words is obviously relevant cross-linguistically, and before coming back to *plus* and its place within the system of French n-words, it is useful to briefly consider the more general class of n-words as it has been defined cross-linguistically.

3.2 Giannakidou's definition

According to Giannakidou (2007), the term n-word was coined by Laka Mugarza (1990) to refer to nominal and adverbial phrases that appear in NC structures. Cross-linguistically the class is wide and heterogeneous, and it is better to use distributional criteria to define it, rather than semantic features. Here is Giannakidou's definition (slightly revised).

An expression α is an n-word if and only if:

- a. α can be used in structures containing sentential negation or another n-word yielding a reading equivalent to one logical negation; and
 b. α can provide a negative fragment answer.

This definition is not very appropriate for French, because it treats on a par the combination of an n-word with sentential negation (considered impossible in French) and the combination of an n-word with another n-word. However, thanks to the disjunction in the definition, it is correct to say that French n-words fit with the first part of the definition. Most French n-words, in addition, also fit with the second part of the definition: *rien*, *personne*, *nullement* can be used in isolation to answer a question and then give only rise to a negative interpretation. Thus the class of French n-words would seem to be defined by two properties:

- a. They can combine and give rise to a negative concord reading;
- b. They can provide a negative fragment answer.

Does *plus* verify these two properties? It can combine with other n-words and give rise to negative concord. But it does not satisfy the requirements for the second property: it cannot be used alone as a fragment answer (17).

- (17) A. *Est-ce que Jean est absent ?*
 “Is Jean absent?”
 B. #*Plus*.
 n-more (Intended meaning: “not anymore”)

Let us now have a closer look at the semantic and distributional properties of *plus* to get a better view of what separates it from the other n-words on one hand and from the sentential negation on the other.

4. Specificities of *plus*

4.1 *Plus is not quantificational*

Most authors merge more or less implicitly the class of n-words and the class of negative quantifiers (see Corblin & Tovena 2003, for instance). Furthermore the most representative n-words in French, be they indefinite pronouns (*rien*, *personne*), determiners (*aucun*), or adverbs (*nulle part*, *jamais*) semantically combine a negation and an existential quantification. Across syntactic categories, the compositional contribution of these words (when they are interpreted negatively) shows a remarkable regularity, the only variable aspect being the domain of quantification (objects, people, locations, times, etc.):⁵

5. We chose the ‘negative existential’ version: $\neg\exists x (Rx \wedge Sx)$ of these formulae, which are logically equivalent to a ‘universal negative’ version $\forall x (Rx \rightarrow Sx)$. This equivalence makes it irrelevant to distinguish the two formulae, even though this is sometimes done in the literature – see for

- (18) *rien* “nothing”: $\lambda W. \neg \exists x (Ox \wedge Wx)$
personne “nobody”: $\lambda W. \neg \exists x (Px \wedge Wx)$
nulle part “nowhere”: $\lambda W. \neg \exists x (Lx \wedge Wx)$
jamais “never”: $\lambda W. \neg \exists x (Tx \wedge Wx)$
aucun (N) “no (N)”: $\lambda N \lambda W. \neg \exists x (Nx \wedge Wx)$

The n-word *plus* does not correspond to this pattern: its asserted contribution is reduced to a sentential negation. As for its presupposed contribution, it cannot be put into a form similar to what we have in (18) (see § 4.3).

Thus, *plus* has to be distinguished from the most central and canonical subset of n-words (across languages) which all share this property of being quantificational – see for instance the list of n-words for Romance languages proposed by Corblin & Tovenia (2003). From this perspective, *plus*, along with a number of other negative items (such as *sans* “without”), needs a place of its own in the system of negation in French.

4.2 Plus+n-word as phrases

Plus can occupy two positions when it co-occurs with another n-word in a sentence. It can either occur immediately after the finite verb, in the canonical position of the sentential negation *pas*, or appear just before the n-word so that together they form a single phrase. This positional property is very visible when *plus* combines with *rien* or *personne* which can easily occupy a preverbal position (19b). It is more difficult to show with the negative adverbs *nulle part* and *jamais* whose canonical position is postverbal. It clearly emerges however when the adverbs are in a preverbal position, as in (20).

- (19) a. *Personne ne dort plus.*
 n-body NE sleeps n-more
 b. *Plus personne ne dort.*
 n-more n-body NE sleeps
 “Nobody sleeps anymore.” (preferred reading)
- (20) a. *Après cela, jamais Jean n’ est plus venu.*
 after that, n-ever Jean NE is n-more come
 b. *Après cela, plus jamais Jean n’ est venu.*
 after that, n-more n-ever Jean NE is come
 “After that, Jean never came anymore.”

instance Giannakidou (2007: 332) who draws from the fact that two logical representations exist (dubbed ‘universal negation’ vs ‘existential negation’) the hypothesis that “some n-words would correspond to existential quantifiers under negation, some others to universal quantifiers, and some others perhaps to both”.

The (a) and (b) sentences are not strictly equivalent. The (a) sentences are semantically ambiguous and can receive either a NC or a DN interpretation. On the contrary, the (b) sentences can be interpreted only as NC instances. For example, (19b) would not be appropriate in the context of (21) which induces a double negation reading. In fact *plus* can combine in this way with many n-words (*rien/personne/nulle part/aucun...*), always giving rise only to a double negation reading.

- (21) A: *Il faudrait être trois. Qui pourrait nous aider? Qui ne dort plus ?*
 “We need to be three. Who could help us? Who is no longer sleeping?”
 B: a. *Personne ne dort plus. Il est encore très tôt.*
 b. *#Plus personne ne dort. Il est encore très tôt.*
 “Nobody sleeps no more. It is still very early.”

We propose to associate the (a) versions, which have two distinct readings, with two distinct syntactic analyses. In the case of the NC reading, we assume that *plus* is not negative. We analyse it as a modifier of the quantificational n-word, which is moved on the surface, exactly as do with other floating adverbs such as *tous* “all” (22) or *beaucoup* “many” (23).

- (22) a. *Tous les enfants sont malades.*
 all the children are sick
 b. *Les enfants sont tous malades.*
 the children are all sick
 “All the children are sick.”
 (23) a. *Jean a lu beaucoup de livres.*
 Jean has read many of books
 b. *Jean a beaucoup lu de livres.*
 Jean has many read of books
 “Jean has read many books.”

In the case of a DN reading, on the contrary, *plus* is indeed negative. It occupies the same position as the one *pas* would occupy in a simple negative sentence, and it introduces a presuppositional content in addition to its negative asserted content. Since it combines with a sentence already bearing a negation induced by an n-word, we obtain a DN reading.

Let us now consider sentences where *plus* occurs in a non-verbal position. We claim that in this configuration *plus* occupies the same syntactic position as other adverbs such as *presque* “almost” or *absolument* “absolutely” and that it forms a phrase with the n-word.

- (24) a. *Presque personne n' a réussi.*
 almost n-body NE has succeeded.
 "Almost nobody succeeded."
 b. *Absolument rien ne change.*
 absolutely n-thing NE changes
 "Absolutely nothing changes."

Nevertheless, there are important differences between *plus* and these adverbs. Depending on whether they occur in a pre- or postverbal position, the meaning of the whole sentence changes radically. When the adverb scopes on the n-word, it changes the meaning of the quantifier, and when it scopes on the verb, it modifies the verb meaning. *Almost nobody* means *somebody* while *almost succeed* means *fail*.

- (25) a. *Personne n' a presque réussi.*
 n-body NE has almost succeeded.
 "Nobody has almost succeeded."
 b. *Rien ne change absolument.*
 n-thing NE change absolutely
 "Nothing changes absolutely."

Nothing similar happens with *plus*. In the case of the NC reading, the position of *plus* has no effect on the assertive content. It only has an effect on the presuppositional content: when *plus* occurs in preverbal position, the presupposition is existential, while when it occurs in postverbal position, the presupposition is underspecified and can be either universal or existential, depending on the speakers. We will consider this in greater details in § 4.3.

At this point, we distinguish two uses of *plus*.

- An instance of *plus* which modifies an n-word, and can appear on its left, as in (19b), or adjoined to the finite verb, as in (19a). It is not negative in the sense that it can only give rise to a NC reading.
- An instance of *plus* which is the presuppositional variant of the sentential negation *pas* and occurs in the adverbial position where *pas* occurs. Like *pas*, this instance of *plus* can also be used in phrasal negations as in (26) (one typical property of phrasal negation in French is that it cannot occur with the particle *ne*).

- (26) a. *Jean a bu du vin (pas / plus) très bon.*
 Jean has drunk of wine (not / n-more) very good
 b. **Jean n' a bu du vin (pas / plus) très bon.*
 Jean NE has drunk of wine (not / n-more) very good
 "Jean drank wine that is not very good (any more)."

It is important to observe that these two uses can coexist in the same sentence, in which case the double negation reading becomes the preferred reading (27).

- (27) *Grâce à ce médecin, Jean n' a plus vu plus personne.*
 thanks to this doctor Jean NE has n-more seen n-more n-body
 "Thanks to this doctor, Jean has no longer seen no one anymore."

On the contrary, (28), in which several quantificational n-words co-occur, is ambiguous, giving rise both to a NC and a DN reading, with a preference for the NC reading. The NC reading is also preferred in (29), where there are as many occurrences of *plus* as occurrences of other n-words.

- (28) *Personne n' a jamais rien dit à personne.*
 n-body NE has n-ever n-thing said to n-body
 "No one has (n)ever said anything to anybody."
- (29) *Plus personne ne dort plus nulle part.*
 n-more n-body NE sleeps n-more n-where
 "No one sleeps (no/any)where anymore."

4.3 *Plus is presuppositional*

The combination *ne ... plus* is usually considered as a presupposition trigger: an utterance such as (30) includes contents of two different types: an asserted content (31a), and a presupposed content (31b).

- (30) *Jean ne dort plus.*
 Jean NE sleeps n-more.
 "Jean doesn't sleep any more."
- (31) a. Assertion: Jean is not sleeping now.
 b. Presupposition: Jean was sleeping (before).

Let us now consider what happens when the trigger *ne ... plus* interacts with a negative quantified expression, for instance the n-words *rien*, or *personne*.

Before going into more detail, it is worth recalling that independently of the case of *plus*, the question of the right presupposition when a trigger is in the scope of a universal quantifier has not received a satisfactory answer in the literature, and both the empirical and the theoretical sides of the question are still the subject of a lively debate (Chemla 2009; Sudo et al. 2012; Zehr et al. 2016). To briefly summarize the situation, a sentence such as (32) may be interpreted, as far as presupposition is concerned, in three different ways (32a–c). In this example, the verb *to win* is a presupposition trigger, and presupposes a participation.

- (32) *None of the bears won the race.* (Zehr et al. 2016, Example (5))
- a. existential: At least one of the bears participated.
 - b. universal: All of the bears participated.
 - c. presuppositionless:⁶ No commitment as to whether some, all or none of the bears participated.

Empirically, there is no consensus on which of these options prevails in general, and indeed several experimental studies have shown variation among speakers: some seem to prefer universal readings, while others seemingly never consider universal presupposition (Sudo et al. 2012). It has not been established yet how consistent speakers are in fact, and what contextual factors influence their interpretation. On the theoretical side, theories of presupposition projection in quantificational sentences can be roughly divided into three groups: theories that only predict universal projection, theories that only predict existential projection, and theories that allow for both projection options (Zehr et al. 2016: 757). In any case, the accepted view is that quantified sentences of the form illustrated in (33) (of which (32) is an instance) give rise to quantified presuppositions, independently of the specific quantifier Q. These presuppositions may be either existential (33a) or universal (33b), and they cannot be trivially predicted compositionally: appropriate projection rules have to be postulated.

- (33) $[Qx : R(x)] S_{p(x)}(x)$ ⁷
- a. $[\exists x : R(x)] p(x)$
 - b. $[\forall x : R(x)] p(x)$

Let us now return to the case of *plus*. We have seen in § 4.2 that it can occupy two distinct positions, and since in one of these positions there is an ambiguity between two readings, we have three different cases to consider. When *plus* is in a postverbal position, we obtain either a DN or a NC reading (34a–b); when *plus* is preposed, it can only give rise to a NC reading (34c).

- (34) a. *Personne ne danse plus.* DN – “Everybody is still dancing.”
 b. *Personne ne danse plus.* NC – “Nobody dances anymore”
 c. *Plus personne ne danse.* NC – “Nobody dances anymore” [*lit.* “n-more n-body”]

6. Zehr et al. (2016) unfortunately call ‘presuppositionless’ a situation in which there is in fact the presupposition that none of the bears participated. We do not consider the case of a lack of presupposition in the rest of the paper.

7. This notation is taken from Chemla (2009). Q is a quantifier, R its restrictor, S its scope, and p a presupposition triggered inside S.

Let us start with the DN reading (34a). We expect *plus*, as a presupposition trigger, to behave similarly to the other triggers: we do in fact obtain either a universal or an existential presupposition depending on the speaker. This variation is exactly the same with (35), the positive equivalent of (34a): the presupposition trigger *encore* is under the scope of a positive universal quantifier, and it triggers an underspecified quantified presupposition *somebody danced* or *everybody danced*.

- (35) *Tout le monde danse encore.*
 “Everybody is still dancing.”

As for NC readings, the situation is slightly less straightforward: out of context, it is quite easy for (34b) and (34c) to lead to either an existential or a universal presupposition reading.

It is however possible to find examples where a stronger preference seems to be at stake. For instance, (36) is a case where the most plausible presupposition is universal (all men were bachelors at some point). Our native speaker impression about this example is that the version with preposed *plus* is degraded, precisely because the existential presupposition that would be preferred with this word order is pragmatically infelicitous.

- (36) a. *Aucun homme marié n' est plus célibataire.*
 n-no man married NE is n-more bachelor
 b. *#Plus aucun homme marié n' est célibataire.*
 n-plus n-no man married NE is bachelor
 “No married man is a bachelor any more.”

On the contrary, at least for some speakers, it seems that when the universal presupposition is highly implausible, as in (37), the version with *plus* in the postverbal position is slightly degraded.

- (37) a. *?Personne ne lit plus Proust.*
 n-body NE reads n-more Proust
 b. *Plus personne ne lit Proust.*
 n-plus n-body NE reads Proust
 “Nobody reads Proust any more.”

At this stage, we can say that the two possible positions of *plus* each come with a preferred presupposition: preverbal *plus* prefers an existential presupposition, while postverbal *plus* prefers the universal presupposition. Our claim is that the preference for an existential presupposition with preverbal *plus* is predicted by the fact that the presupposition has a wide scope over the quantified negative sentence, while the preference (for some speakers) for a universal presupposition in

postverbal position is not semantic but pragmatic and comes from the competition between the two positions. In the next section we spell out the compositional analysis that we hypothesize.

5. Analysis

We summarize in Figure 1 the various readings obtained when *plus* interacts with a quantificational n-word like *personne*. When *plus* is in postverbal position, the two readings we obtain differ on the asserted dimension (DN vs NC) but not on the presupposed dimension: in both cases we obtain an underspecified presupposition that can be either universal or existential. When *plus* forms a phrase with an n-word, then we obtain only a NC reading (labelled NC₂), which has the same asserted content as the other variant of NC reading (labelled NC₁), but gives rise only to the existential presupposition.

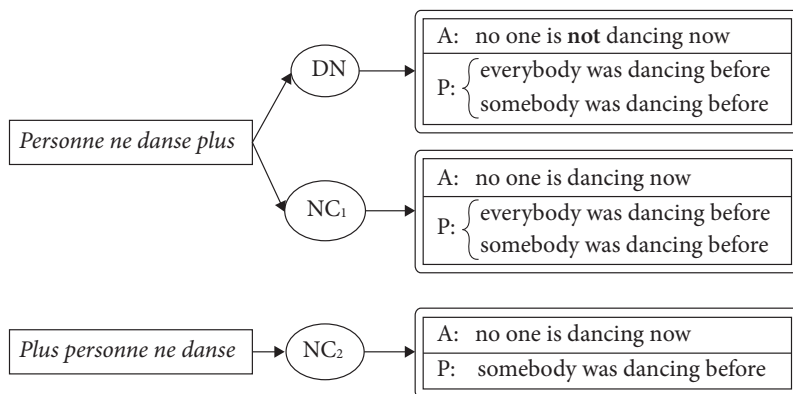


Figure 1. Possible readings for (34)

We have proposed in § 4.2 associating different syntactic structures to sentences built with *plus* and an n-word. We have assumed that when *plus* occurs just before a quantificational n-word, it forms a phrase with it. It is analysed as a presuppositional adverb, which adds only a presuppositional content to the rest of the sentence. When *plus* occurs in the postverbal position, the sentence is syntactically ambiguous:

- Either *plus* bears on the VP, which is dominated by a NegP, exactly as in the case of utterances built with the sentential negation *pas*. The only difference between *pas* and *plus* is that in the case of *plus*, a presupposition is conveyed in addition to this negative content. Since the sentence involves an n-word and a plain negation, it gives rise to the DN reading.

- Or *plus* occupies a postverbal position, but is indeed a floating adverb: it forms a phrase with the n-word, even if it moves at the surface level.

Let us now try to predict these various readings, on both the assertion and presupposition sides. To represent the semantic contributions of *plus* in the different cases, we introduce two semantic operators, labelled plus_p and plus_s . The letters p and s are used to refer respectively to the phrasal and the sentential versions of *plus*. These operators are defined on the two assertive (A) and presuppositional (P) dimensions in the following way: they bring the same content on the presupposed level (the negation of the fact that the argument of *plus* was the case in the past), but they either make no contribution on the asserted level (in technical terms, it is associated with the identity function) or introduce a negation (38).

- (38) plus_p A: $\lambda P. P$
 P: $\lambda P. \exists t (t < n \wedge \neg P(t))$
 plus_s A: $\lambda P. \neg P$
 P: $\lambda P. \exists t (t < n \wedge \neg P(t))$

The intuition behind this proposal is that the presuppositional contribution of *plus* (whether it is a full negative adverb or concord n-word) can be seen as operating on a description of the present state-of-affairs, stating that the present state-of-affairs was not holding at least at one temporal point in the past.⁸

Let us now see how this proposal makes it possible to account for the various interpretations that were observed.

In the simplest case, where *plus* occurs without quantified n-words, we analyse *plus* as a sentential negation. It scopes over the positive proposition corresponding to the sentence without *plus*. Semantically, it contributes the operator plus_s , which conveys the same assertive content as *pas*, and in addition a presuppositional content:⁹

- (39) a. *Jean ne danse plus.*
 “Jean dances no more.”
 b. plus_s (*Jean danse.*)
 c. A: Jean is not dancing now.
 P: Jean used to dance before.

8. Technically, this requires that the state-of-affairs be represented by functions from time points to truth values.

9. Diachronically, the modern form *ne ... plus* comes from a time when *ne* was strong enough to express sentential negation, so that the old French version of *Jean ne danse plus* was to be analysed exactly as modern *Jean ne danse pas plus*.

In the case where *plus* is preverbal and modifies the n-word, it corresponds to the operator plus_p . It does not introduce a negation, but only a presupposition that combines with the assertive content conveyed by the rest of the sentence (which is negative thanks to the other n-word) (40). The asserted part is directly computed compositionally (the identity function applying on the semantic content of *nobody dances now*), and the presupposed part is an existential presupposition, obtained through a computation that could be paraphrased as in (41).

- (40) a. *Plus personne ne danse.*
 b. plus_p (personne_x (x danse))
 c. A: Nobody is dancing now.
 P: Somebody used to dance before.
- (41) There is a time t before ‘now’ where it was not the case that no one is dancing at t .

In the two remaining cases, *plus* is in a postverbal position, and falls within the scope of a negative quantifier. We can then predict an underspecified presupposition (existential or universal) just as in the general case of interaction between quantification and presupposition. The DN and NC_1 cases differ only on the asserted content, which is obtained compositionally by making the following assumptions. In the case of a DN reading, we assume that the word *plus* corresponds to the sentential negation plus_s and we derive the right content (42). In the case of an NC reading, we assume that *plus* contributes the operator plus_p , and in this way we can also derive compositionally the correct asserted content (43).

- (42) a. *Personne ne danse plus.* (DN)
 b. personne_x (plus_s (x danse))
 c. A: Nobody is **not** dancing now.
 P: { Everybody was dancing before.
 Somebody was dancing before.
- (43) a. *Personne ne danse plus.* (NC)
 b. personne_x (plus_p (x danse))
 c. A: Nobody is dancing now.
 P: { Everybody was dancing before.
 Somebody was dancing before.

So the ambiguity of sentences combining the postverbal *plus* and an n-word comes from two different syntactic forms, as illustrated in (42) and (43). In both cases the quantifier scopes over *plus*, but in one case, it is interpreted as plus_s and in the other case as plus_p . At this point it is useful to consider the syntax of these various configurations. We can adopt Zeijlstra’s (2014) analysis of French negation. We just have

to add that there are two different *plus* in French. The first one corresponds to plus_s ; it would be analysed syntactically in a similar way to *pas*, and it gives rise to a DN reading when combined with n-words. The second one (plus_p) is not an n-word, but a simple floating adverb, which only adds a presupposition to the assertive content conveyed by the sentence it combines with. There is no need to include it in the set of n-words: since it does not introduce any variable in itself, it has no reason to participate in a system of unselective binding by a negative operator.

6. Conclusion

In this work, we have shown that *plus* turns out to be unique in the system of negation in French. On the one hand, *plus* is comparable to the negative adverb *pas*, since it contributes a negation to the assertive content of the sentence. But unlike *pas*, it can co-occur with n-words. On the other hand, *plus* is similar to n-words, with which it can combine, yielding a concord reading. But unlike most n-words, it is not quantificational, and it can form a phrase with an n-word, which it then modifies by adding only some presuppositional content. This leads us to conclude that *plus* contributes two different operators, which however share a central property: both trigger the same presupposition.

Our proposal that the operator plus_s introduces both a sentential negation and a presupposition is in line with what is suggested by diachrony: at one point in history, the combination *ne ... plus* was composed of the sentential negation *ne* and a comparative adverb only in charge of the temporal presupposition. But while the weakening of *ne* as a full negation led to *pas* becoming the main term for sentential negation, the adverb *plus* partly took over the role of *pas* so that instead of the combination *ne ... pas plus* (which corresponds to what we find in other languages: e.g., It. *non più*, Ge. *nicht mehr*) we ended up with the present version *ne ... plus*. This fact also explains why it is impossible to find in contemporary French a combination of *plus* with the negative adverb *pas*, whereas, even if this type of combination remains quite rare, it is possible for the other n-words.

Acknowledgments

The authors wish to thank two anonymous reviewers, as well as the audiences of the International Conferences “The Pragmatics of Negation – Aspects of Communication” (Stockholm, 2017) and “Going Romance” (Bucharest, 2017). Shirley Carter-Thomas provided us with invaluable help in preparing the final version of this paper.

References

- Chemla, Emmanuel. 2009. Presuppositions of quantified sentences: Experimental data. *Natural Language Semantics* 17(4). 299–340. <https://doi.org/10.1007/s11050-009-9043-9>
- Corblin, Francis, Viviane Déprez, Henriette de Swart & Lucia Tovena. 2004. Negative Concord. In Francis Corblin & Henriette de Swart (eds.), *Handbook of French semantics*, 417–452. Stanford: CSLI Publications.
- Corblin, Francis & Lucia Tovena. 2003. L'expression de la négation dans les langues romanes. In Daniele Godard (ed.), *Les langues romanes: Problèmes de la phrase simple*, 281–343. Paris: CNRS Éditions.
- De Swart, Henriette & Ivan A. Sag. 2002. Negation and negative concord in Romance. *Linguistics and Philosophy* 25(4). 37–417. <https://doi.org/10.1023/A:1020823106639>
- Giannakidou, Anatasia. 2007. N-words and negative concord. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, 327–391. Malden, MA: Blackwell.
- Laka Mugarza, Miren Itziar. 1990. Negation in syntax: On the nature of functional categories and projections. Doctoral dissertation, MIT.
- Larrivé, Pierre. 2016. The pragmatics of marked configurations: Negative doubling in French. *Journal of Pragmatics* 95. 34–49. <https://doi.org/10.1016/j.pragma.2016.02.002>
- Muller, Claude. 1991. *La négation en français*. Geneva: Droz.
- Sudo, Yasutada, Jacopo Romoli, Martin Hackl & Danny Fox. 2012. Presupposition projection out of quantified sentences: Strengthening, local accommodation and inter-speaker variation. In Maria Aloni, Vadim Kimmelman, Floris Roelofsen, Galit W. Sassoon, Katrin Schulz & Matthijs Westera (eds.), *Logic, language and meaning*, 210–219. Berlin: Springer. https://doi.org/10.1007/978-3-642-31482-7_22
- Zehr, Jérémy, Bill Cory, Tieu Lyn, Jacopo Romoli & Florian Schwarz. 2016. Presupposition projection from the scope of none: Universal, existential, or both?. In Mary Moroney, Carol-Rose Little, Jacob Collard & Dan Burgdorf (eds.), *Proceedings of the 26th Semantics and Linguistic Theory Conference (SALT 26)*, 754–774. LSA and CLC Publications. <https://doi.org/10.3765/salt.v26i0.3837>
- Zeijlstra, Hedde. 2009. On French negation. In Iksoo Kwon, Hannah Pritchett & Justin Spence (eds.), *Annual Meeting of the Berkeley Linguistics Society* 35, 447–458. Berkeley: Berkeley Linguistic Society. <https://doi.org/10.3765/bls.v35i1.3631>
- Zeijlstra, Hedde. 2014. On the uninterpretability of interpretable features. In Peter Kosta, Steven L. Franks, Teodora Radeva-Bork & Lilia Schürcks (eds.), *Minimalism and beyond*, 109–129. Amsterdam: John Benjamins. <https://doi.org/10.1075/lfab.11.05zei>

An experimental approach to parallelism in ellipsis

The case of *pro*-drop in Romance gapping

Gabriela Bilbiie and Israel de la Fuente

University of Bucharest / University of Lille

The present study investigates parallelism constraints in Romance gapping with *pro*-drop. We provide empirical evidence from two acceptability judgment tasks in Romanian and Spanish, that syntactic parallelism in gapping is less strict than traditionally assumed. Parallelism constraints are stronger at the discourse and information structure levels than at the syntactic level. Our experimental findings show that gapping with *pro*-drop is as acceptable as without *pro*-drop, challenging the double foci analysis and supporting the topic-focus pattern for gapping constructions. At the semantic level, our data show that the contrast in gapping is not always established explicitly; more generally, they show that one cannot make a correlation between contrast and subject (null vs overt) expression in Romance.

Keywords: gapping, parallelism, *pro*-drop, constructional approach, Romance, Romanian, Spanish, acceptability judgments

1. Introduction

A central topic in the literature on ellipsis is, on the one hand, the identity between the missing material and its antecedent, and, on the other hand, the parallelism between remnants and correlates. In this chapter, we focus on parallelism constraints in gapping and, in particular, on the most studied type of parallelism, involving syntactic identity. Gapping, illustrated in (1), refers to an elliptical clause (called the target or the gapped clause) containing at least two lexically-realized elements (= remnants, e.g., *Leslie* and *German*, one of them being generally the subject) and missing at least the head verb (= the gap); this elliptical clause is combined with a complete clause (= the source) which provides the necessary material for its interpretation: the antecedent of the gap (e.g., *speaks*) and the elements which are parallel to remnants (= correlates, e.g., *Robin* and *French*).

(1) *Robin speaks French [and Leslie German].*¹

Since Ross (1967, 1970), gapping has received a lot of attention in English, German (Hartmann 2000; Winkler 2005; Repp 2009, among others), and other languages (in particular, Japanese and Korean). However, gapping has by and large been overlooked in Romance, though see Abeillé et al. 2014 and Bilbiie 2017 for an exhaustive analysis of gapping in French and Romanian, and also Centeno 2011 for Spanish, Silva 2014 for European Portuguese and Dagnac 2016 for French. The present study investigates gapping in Romanian and Spanish. A major advantage of studying gapping in Romance concerns the fact that Romance languages allow us to better observe the behavior of gapping, in particular with respect to parallelism constraints. They allow us to test more ‘connectivity effects’ between the gapped clause and the source, in terms of syntactic category, word order or number of constituents. More specifically, Romance languages such as Romanian, Italian or Spanish provide us with the perfect ground to assess the interaction between gapping and *pro*-drop phenomena.

The main goal of our study is to demonstrate on experimental grounds that syntactic parallelism in gapping is less strict than traditionally assumed. This result has some major implications for the syntactic analysis, information structure and semantics of gapping. On a methodological level, the main goal of this contribution is to go beyond the limited data paradigms obtained through introspection, by showing how psycholinguistic experiments can be used to obtain a much finer-grained perspective on data pertaining to ellipsis phenomena.

The chapter is structured as follows: Section 2 discusses the syntactic parallelism between remnants and correlates. Section 3 details two acceptability judgment tasks testing the case of *pro*-drop in Romanian and Spanish gapping. Section 4 proposes a construction-based analysis of gapping compatible with the experimental findings. Section 5 discusses some implications of our results for the information structure and semantics of gapping.

2. Syntactic parallelism in gapping

The standard assumption for gapping is that a structural parallelism should hold between the target clause and its source (e.g., Sag 1976; Hartmann 2000). Thus, according to Hartmann (2000: 162), remnants “must be syntactically identical” to their correlates. She defines this structural parallelism as an identity on the syntactic form and function between a remnant and its correlate: on the one hand, remnants

1. The material serving as antecedent for the gap is underlined throughout this chapter.

and correlates must be of the same syntactic category (condition on categorial identity), and, on the other hand, remnants and correlates must have the same grammatical function (condition on grammatical function identity). This idea is captured by the so-called ‘connectivity effects’ (see Merchant 2001, 2004): remnants and correlates share the same case or preposition marking, the same syntactic category, same number of remnants and correlates, same word order, etc., offering a strong argument for structural approaches in terms of syntactic reconstruction.

This structural parallelism is considered to mainly contribute to the resolution of ellipsis. The same idea occurs in previous work on parallelism in the processing literature (Frazier et al. 1984, 2000; Carlson 2001, 2002, among others), in which it is shown that “parallelism of many types is helpful to the processor, in that the second conjunct is easier to process if it is parallel to the first” (Carlson 2002: 4).

The same parallelism constraints were widely discussed in relation to coordination in general, as shown by the so-called Coordination of Likes Constraint (Chomsky 1957; Schachter 1977; Jackendoff 1977; Williams 1978, and subsequent literature). Therefore, according to Chomsky’s (1957: 36) generalization, syntactically different categories cannot be conjoined. As gapping mostly occurs in coordination structures, it is not surprising at all that the same parallelism constraints were applied to gapping constructions.

The assumption requiring a strong syntactic parallelism in coordination and, consequently, in gapping structures was first challenged by Sag et al. (1985) and Grosu (1985, 1987). Grosu offers some intuitive evidence that syntactically non-parallel constituents (in terms of syntactic category or syntactic function) could be acceptable in gapping (2), under certain pragmatic and focus conditions.

- (2) a. *John can drink ANYTHING, and Mary, with ANYBODY.*
(Grosu 1987: 451 Example (51c))²
- b. *John drinks only the BEST wines, and Mary, only in the BEST company.*
(Grosu 1987: 451 Example (51a))
- c. *John sings the MOST inappropriate songs, and Mary – at the MOST inappropriate hours.*
(Grosu 1985: 234 Example (11a))
- d. *John has written FIFTEEN articles, and Mary – to TWO HUNDRED subscribers.*
(Grosu 1985: 234 Example (11b))

Sag et al. (1985) convincingly show that the Coordination of Likes Constraint is too strong for coordination in general and for gapping in particular. First, the order of remnants in the gapped clause does not necessarily need to parallel that of their correlates in the source clause (3), provided that this order is licit in the grammar. Second, remnants may differ from their correlates with respect to their syntactic

2. The small capitals mark prosodic stress.

category (4a), provided that each remnant matches some subcategorization frame of the verbal predicate in the source. We can thus explain the contrast between (4a) and (4b): unlike the AP *crazy* or the NP *an incredible bore*, the PP *in good spirits* does not match the subcategorizational requirements of the verb *become*, as illustrated in (4c).

- (3) *A policeman walked in at 11, and at 12, a fireman.*
 (Sag et al. 1985: 158 Example (106))
- (4) a. *Pat has become [crazy]_{AP}, and Chris [an incredible bore]_{NP}.*
 (Sag et al. 1985: 160 Example (113b))
- b. **Pat has become [crazy]_{AP}, and Chris [in good spirits]_{PP}.*
 (Sag et al. 1985: 160 Example (113c))
- c. *Pat has become {crazy | an incredible bore | *in good spirits}.*

The conclusion that can be drawn from these examples is that remnants and their correlates in gapping constructions obey the same syntactic constraint as conjuncts in ordinary constituent coordinations: each must match some subcategorization of the shared predicative material, in accordance with Wasow's Generalization, governing the so-called Coordination of Unlikes. We recall here the formal statement of this generalization as given by Pullum & Zwicky (1986: 752–753): "If a coordinate structure occurs in some position in a syntactic representation, each of its conjuncts must have syntactic feature values that would allow it individually to occur in that position". This accounts for the grammaticality of (5a), where the AP *crazy* and the NP *an incredible bore* are individually compatible with the verb *become*; at the same time, it rules out (5b), since the verb *become* cannot be construable with the PP *in good spirits*.

- (5) a. *He has become [crazy]_{AP} and [an incredible bore]_{NP}.*
 b. **He has become [crazy]_{AP} and [in good spirits]_{PP}.*

As shown by Abeillé et al. (2014) and Bilbiie (2017), Romance languages parallel English in this respect: order and category asymmetries are allowed, provided the syntactic constraint mentioned above is observed.

However, syntactic asymmetries related to the number of remnants and correlates in gapping have never been investigated in detail or empirically tested before. As Romance languages such as Romanian, Italian or Spanish are *pro*-drop languages, they allow us to better observe this kind of syntactic asymmetry in gapping. Based on some linguistic intuitions, Abeillé et al. (2014: 251) hypothesize that constructed examples such as (6) in Romanian or (7) in Italian, where there is no overt subject correlate in the source, should be perfectly acceptable, suggesting therefore that the number of remnants may differ from the number of lexically-realized correlates.

- (6) *Lunea merg la film, iar sora mea la muzeu.*
 Monday.DEF go.PRS.1SG to film and sister POSS.1SG to museum
 “On Monday, I go to the cinema, and my sister (goes) to the museum.”
- (7) *Mangio la pasta e Giovanni il riso.*
 eat.PRS.1SG DEF pasta and Giovanni DEF rice
 “I eat pasta and Giovanni (eats) rice.”

Our general hypothesis is that syntactic parallelism in gapping does not operate at the level of phrase structure; as there is no strict syntactic parallelism in gapping, one expects that examples such as (6) and (7) would be perfectly acceptable. Therefore, we expect no significant difference between the acceptability of these gapping examples with *pro*-drop and their counterparts without *pro*-drop.

3. Experimental studies on *pro*-drop and gapping

The present study was motivated by the lack of empirical data on gapping with *pro*-drop, and by the fact that such structures (see (6)–(7) above) should be ruled out according to a strict syntactic parallelism approach of ellipsis.

In what follows, we report results from two acceptability judgment tasks (AJTs), one in Romanian and one in Spanish, which constitute evidence in favor of the hypothesis that syntactic parallelism in gapping is less strict than has been traditionally assumed.

Participants

A total of 68 Romanian native speakers (mean age: 25.3; range: 18–48) and 65 Spanish native speakers (mean age: 30.7; range: 15–55) completed the AJTs. Participants were all from Romania and Spain, respectively.

Materials

We created 24 experimental items following a 2×2 factorial design with GAPPING (+Gapping, –Gapping) and PRO-DROP (+*pro*-drop, –*pro*-drop) as independent variables. This manipulation gave rise to the 4 experimental conditions shown in (8) and (9), for Romanian and Spanish, respectively. We compared elliptical occurrences (conditions a-b) with non-elliptical ones (conditions c-d), in order to better control our two factors and to rule out other explanations for any treatment effects that we might observe. As the examples show, the experimental items were translation equivalents of the same materials in the two languages.

In light of the assumptions presented in Section 2, we expect no significant difference between the conditions (a) and (b). In other words, we expect that gapping with or without *pro*-drop elicit similar acceptability ratings. Furthermore, by comparing a regular gapping construction (condition a) and its non-elliptical counterpart (condition c), we can corroborate previous experimental findings that show cross-linguistic differences with respect to preferences between an elliptical construction and its non-elliptical counterpart. In particular, based on two experimental studies (a written questionnaire and an auditory comprehension study) on English, Carlson (2001) observes a preference for non-gapping over gapping structures. Experimental evidence on embedded gapping in Romance languages does not show a preference for any of these two structures (see Bilbiie & de la Fuente 2019; Bilbiie et al. 2019).

- (8) *Ne-am făcut deja planul pentru weekend.*
 “We have already planned our weekend.”
- a. [+Gapping, *-pro*-drop]
Eu voi merge la film, iar sora mea la muzeu.
 I will.1SG go to film and sister.DEF POSS.1SG to museum
 “I will go to the cinema, and my sister to the museum.”
- b. [+Gapping, *+pro*-drop]
Voi merge la film, iar sora mea la muzeu.
 will.1SG go to film and sister.DEF POSS.1SG to museum
 “(I) will go to the cinema, and my sister to the museum.”
- c. [–Gapping, *-pro*-drop]
Eu voi merge la film, iar sora mea va merge la muzeu.
 I will.1SG go to film and sister.DEF POSS.1SG will.3SG go
 to museum
 “I will go to the cinema, and my sister will go to the museum.”
- d. [–Gapping, *+pro*-drop]
Voi merge la film, iar sora mea va merge la muzeu.
 will.1SG go to film and sister.DEF POSS.1SG will.3SG go
 to museum
 “(I) will go to the cinema, and my sister will go to the museum.”
- (9) *Ya tenemos planes para el fin de semana.*
 “We have already planned our weekend.”
- a. [+Gapping, *-pro*-drop]
Yo iré al cine y mi hermana al museo.
 “I will go to the cinema, and my sister to the museum.”

- b. [+Gapping, +*pro-drop*]
Iré al cine y mi hermana al museo.
 “(I) will go to the cinema, and my sister to the museum.”
- c. [–Gapping, –*pro-drop*]
Yo iré al cine y mi hermana irá al museo.
 “I will go to the cinema, and my sister will go to the museum.”
- d. [–Gapping, +*pro-drop*]
Iré al cine y mi hermana irá al museo.
 “(I) will go to the cinema, and my sister will go to the museum.”

The experimental items were presented following a context sentence, which served as a contextual anchor (cf. Kuno 1976; Steedman 2000). They were all coordinated sentences joined by the coordinative conjunctions *iar*³ “and” in Romanian and *y* “and” in Spanish. The gapped and the source clauses introduced a character by means of a proper name or a definite NP. Special care was taken to avoid potential variability in the experimental items: First, in order to avoid ambiguity, characters were always of a different person (e.g., never 3rd person in both clauses). Second, only verbs that were dynamic and agentive were used (e.g., *to enter, to go out, to arrive, to work, to come, to fly*, etc.). Third, verbs were always in the future tense in order to avoid underspecified verbal forms (in Romanian, the present indicative allows syncretic forms for some persons with some inflectional classes, e.g., third person singular and third person plural, first person singular and third person plural). Finally, the postverbal constituent was always a temporal or locative complement or adjunct (e.g., *to the cinema vs to the museum, at 3pm vs at 5pm*, etc.).

In addition to the 24 experimental items, 20 filler items (from an unrelated experiment) and 16 control items were also included in order to drive the participants’ attention away from the phenomenon under investigation. The fillers featured subject/object relative clauses that contained subject/object-biased implicit causality verbs. The control items also contained relative clauses but this time half were grammatical and half were ungrammatical. As the example in (10) illustrates, the ungrammaticality stemmed from subject-verb agreement errors, which could be potentially overlooked due to attraction effects (i.e., the number of the verb matched the number of the closest NP). Half of the control items were preceded by a context sentence. The control items were used to make sure that participants were completing the task carefully.

3. We used the conjunction *iar* “and” instead of *și* “and”, since this special ‘contrastive’ conjunction is the most used in gapping, being compatible with its specific constraints (e.g., double-contrastiveness, symmetric discourse relations). For more details on *iar* in Romanian, see Bilbiie & Winterstein (2011) and Bilbiie (2017).

- (10) a. *The key that opened the cabinets was in her purse.*
 b. **The mechanic that fixed the wheels were from an Eastern European country.*

Procedure

The experiments were administered on IbexFarm (Drummond 2013). Sentences were presented in a Latin Square within-subject design, so that participants were exposed to all 4 experimental conditions, but never to the same item in more than one condition. After reading the instructions and answering some language-background questions, participants completed 6 practice items, in order to become familiar with the experiment.

Participants were instructed to read the sentences carefully and to judge their acceptability by using a 7-point Likert scale, where 1 means completely unacceptable and 7 means completely acceptable. Participants did not have the option to go back to change previous judgments. The experiment took 10 minutes to complete.

Results

Judgments (1–7) were entered into mixed-effect linear regression analyses, using the lme4 package (Bates et al. 2015) in R (R Development Core Team 2008). Our models included Gapping and *pro*-drop as fixed effects, and Participant and Item as random effects. Uncorrelated slopes of the fixed predictors (i.e., by-subject and by-item random slopes for Gapping**pro*-drop⁴) were entered when justified (Barr et al. 2013).

The mean acceptability judgments for both languages are given in Table 1 and plotted in Figures 1 and 2 for Romanian and Spanish, respectively. As the table shows, judgments for all conditions in both languages were above 6, which indicates that *pro*-drop is acceptable regardless of whether it occurs in a construction with gapping or with a full clause. Although the non-*pro*-drop conditions were judged as (slightly) more acceptable than the *pro*-drop conditions, this difference did not reach statistical significance in either language. The models did not yield any other significant effects or interactions. Summaries of the models' fixed effects are given in Table 2, for Romanian, and 3, for Spanish.

4. Only by-item random slopes for Gapping**pro*-drop were entered in the model used on the Romanian data in order to avoid errors of convergence.

Table 1. Mean acceptability judgments in Romanian and in Spanish (SD in parentheses)

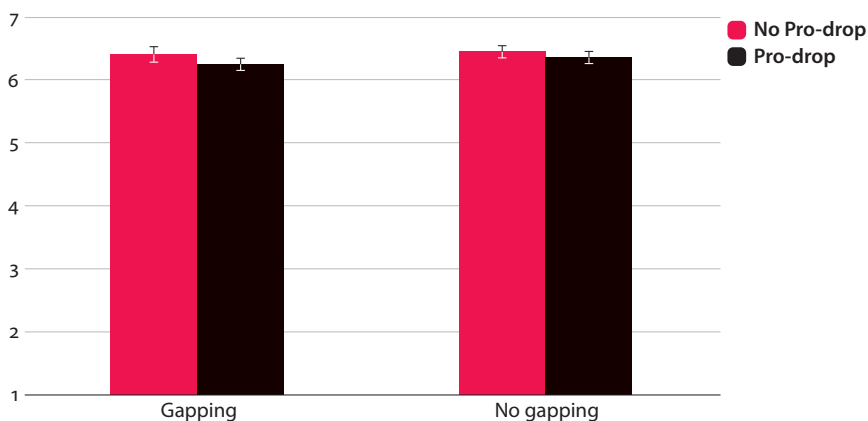
	Romanian		Spanish	
	Gapping	No gapping	Gapping	No gapping
<i>Pro-drop</i>	6.25 (0.90)	6.37 (0.81)	6.35 (0.93)	6.32 (0.97)
No <i>pro-drop</i>	6.43 (0.89)	6.45 (0.82)	6.5 (0.88)	6.55 (0.83)

Table 2. Summary of the model's fixed effects for Romanian

	Estimate	Std. Error	df	t-value	Pr(> t)
(Intercept)	6.42182	0.12119	70.87000	52.991	<.001
Gapping	0.03481	0.07897	77.17000	0.441	0.661
<i>Pro-drop</i>	-0.15873	0.10335	23.96000	-1.536	0.138
Gapping* <i>Pro-drop</i>	0.07455	0.11937	40.17000	0.625	0.536

Table 3. Summary of the model's fixed effects for Spanish

	Estimate	Std. Error	df	t-value	Pr(> t)
(Intercept)	6.50588	0.11382	66.97000	57.158	<.001
Gapping	0.03842	0.10039	54.28000	0.383	0.7034
<i>Pro-drop</i>	-0.15260	0.07945	36.07000	-1.921	0.0627
Gapping* <i>Pro-drop</i>	-0.06701	0.08537	29.85000	-0.785	0.4386

**Figure 1.** Mean acceptability judgments in Romanian

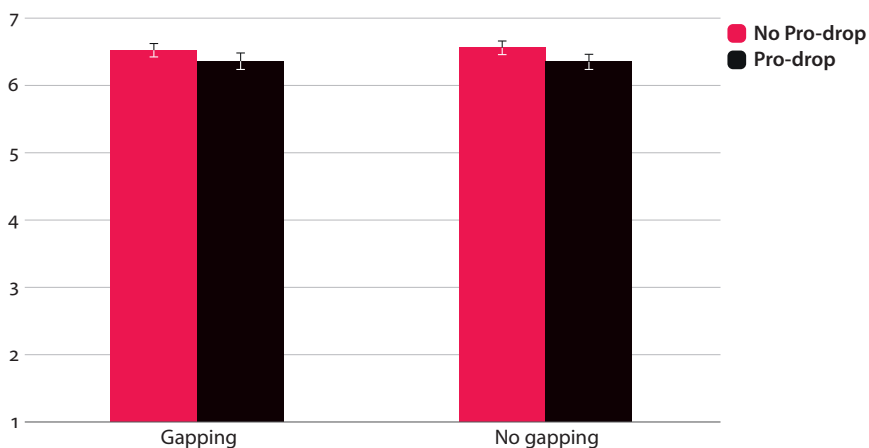


Figure 2. Mean acceptability judgments in Spanish

In order to make sure that participants completed the experiments carefully, we also analyzed their responses to the control items. The models revealed a highly significant effect of grammaticality in both languages (Rom. $X^2(1) = 60.06, p < .001$; Sp. $X^2(1) = 66.273, p < .001$; mean judgments: 6.1 vs 1.9 for Rom.; 6.5 vs 2.1 for Sp.). The mean acceptability judgments for control items are plotted in Figure 3.

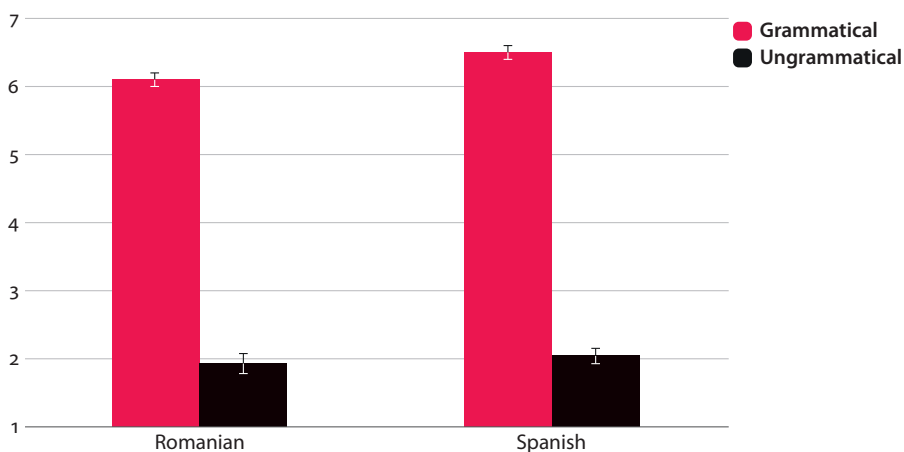


Figure 3. Mean acceptability judgments for control items in Romanian and in Spanish

Discussion

We hypothesized that gapping with *pro*-drop should be acceptable in Romance. This hypothesis is borne out by the results of our experiments, since we observed that gapping with *pro*-drop is as acceptable as regular gapping (judgments for all conditions in both languages were above 6), in stark contrast with ungrammatical control items. These results support the assumption that syntactic parallelism in gapping is not as strong as commonly assumed and, in particular, the fact that the number of remnants may differ from the number of overt correlates. Therefore, the syntactic parallelism constraints proposed by the structural approaches on ellipsis must be reconsidered. We conclude that syntactic parallelism does not operate at the level of phrase structure, but rather at the more abstract level of grammatical functions, as listed in the argument structure of predicates.⁵

Interestingly, our results show a numerical effect of *pro*-drop which is not related to gapping, as *pro*-drop conditions are judged slightly less acceptable than non-*pro*-drop conditions. While the *pro*-drop issue in general is beyond the scope of this chapter, our results call for further research on *pro*-drop in Romance. Specifically, research on the alternation null/overt subjects in different syntactic constructions (cf. Mayol 2010; Filiaci et al. 2014; Biezma 2014) might uncover other factors that constrain this alternation, beyond the traditional factors found in the literature.

As an additional result, the comparison of the conditions (a) and (c), i.e., canonical gapping and its full counterpart, does not support Carlson's (2001) results for English, where there is a preference for non-gapping over gapping. In particular, our results do not show any effect driven by a preference for any of these two structures. This difference between English, on the one hand, and Romanian and Spanish, on the other hand, is confirmed once again by a parallel experiment on

5. In mainstream generative approaches, this claim might seem too strong, since in their perspective the *pro*-drop subject is present at the syntactic level in the form of an empty pronoun (*pro*) and is licensed and interpreted by the inflection of the verb (see Rizzi 1982, 1986). However, in other generative approaches, which avoid unpronounced syntactic material, such as HPSG (Manning & Sag 1999; Ginzburg & Sag 2000), LFG (Bresnan 2001; Bresnan et al. 2015) or Simpler Syntax (Culicover & Jackendoff 2005), the *pro*-drop subject is not represented in the tree structure, but in the lexical argument structure of the verb (see Section 4). Thus, in these approaches, the cases of gapping with *pro*-drop do not display a syntactic parallelism. Moreover, even for approaches which would consider that the verbal inflection is interpretable as a clitic/affix on the verb and this clitic/affix would be the constituent that is paired with the subject of the gapped clause, the syntactic parallelism constraint is problematic, since in that case one would have a lexical subject remnant which has to be paired with a weak element (i.e., an affix). Apart from this problem, one has to note that cross-linguistically, the label '*pro*-drop' is not strictly related to the verbal inflection, since there are *pro*-drop languages, such as Chinese, Japanese or Korean, which do not have inflection to mark subject-verb agreement (see Tomioka 2003: 335).

embedded gapping in these languages, as explained in Bilbîie & de la Fuente (2019) and Bilbîie et al. (2019). More generally, all these results show that languages may differ with respect to the preference between an elliptical construction and its full counterpart.

4. A construction-based analysis of gapping

Three kinds of syntactic analyses have been explored to account for the unusual form/meaning mapping in gapping:

1. The movement-based analysis, rejecting ellipsis, appeals to some leftward movement process that combines across-the-board extraction of the shared head verb out of each conjunct and asymmetric extraction of non-shared constituents preceding the head verb out of the first conjunct (Johnson 2004, 2009). In this account, gapping involves a subclausal coordination. Despite its semantic motivation in gapping (e.g., it accounts for the wide scope of semantic operators such as modals, negation or quantifiers), it faces several empirical problems, as discussed in Abeillé et al. (2014) and Bilbîie (2017): it does not account for complex gaps with non-continuous strings of words; it wrongly predicts the linearization of the correlative (initial) conjunctions in Romance; it does not account for the use of the clausal conjunction *iar* ‘and’ in Romanian.
2. The deletion-based analysis appeals to some deletion process, preceded in some approaches by extraction of remnants in the left periphery (see Ross 1967, 1970; Sag 1976; Hartmann 2000, among others). As it mainly involves syntactic reconstruction of the missing material, it posits a syntactic structure for the gap in the target clause. This analysis still faces several empirical problems, also discussed in Culicover & Jackendoff (2005); Abeillé et al. (2014) and Bilbîie (2017): in particular, it is challenged by the violation of locality constraints in English and Romance, by empirical evidence showing that the target is not a finite clause (e.g., the combination with introducers such as *and not* or *but not* in English, which are incompatible with a finite clause), or by various syntactic mismatches (including the violation of syntactic parallelism constraints discussed in Section 2 of this chapter).
3. The construction-based analysis appeals to a dedicated meaning-form rule, namely a ‘construction’, which maps a headless structure to a clausal meaning (see Sag et al. 1985; Steedman 1990, 2000; Culicover & Jackendoff 2005; Abeillé et al. 2014; Bilbîie 2017, among others). Therefore, it does not involve a syntactic reconstruction of the gap, but rather a semantic reconstruction.

Our data seem to fit better with a construction-based analysis which does not derive the unusual meaning-form mapping in the gapped clause from a hidden syntactic structure (see empirical evidence in favour of this analysis in Abeillé et al. 2014; Bilbiie 2017 and Bilbiie & de la Fuente 2019). This analysis is a non-structural approach, based on a ‘what you see is what you get’ syntactic structure (cf. Ginzburg & Sag’s 2000 theory of fragments or Culicover & Jackendoff’s 2005 Simpler Syntax Hypothesis).

A detailed account of gapping constructions in Romance is developed in Abeillé et al. (2014) and Bilbiie (2017) within a construction-based version of Head-driven Phrase Structure Grammar (HPSG). In this contribution, we simply apply this analysis to our data with gapping and *pro*-drop in Romanian and Spanish. Recall the two samples of our experimental items in (8) and (9) above. A simplified version of the item of interest (with gapping and *pro*-drop) is given in (11) for Romanian. The corresponding syntactic tree is given in Figure 4, which we comment on in detail below.

- (11) *Merg la film, sora mea la muzeu.*
 go.PRS.1SG to film sister.DEF POSS.1SG to museum
 ‘(I) go to the cinema, my sister to the museum.’

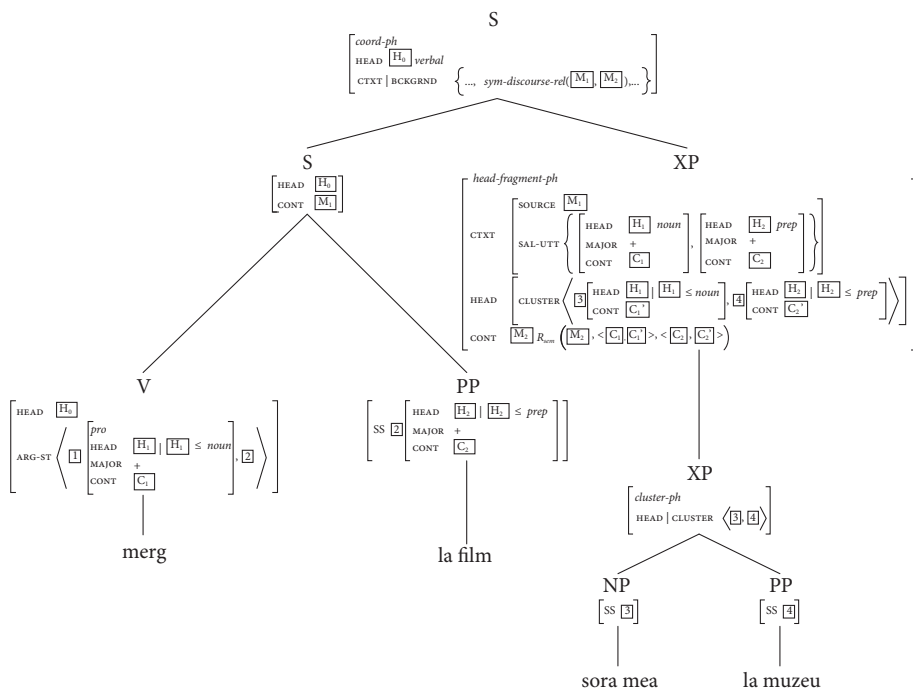


Figure 4. Simplified syntactic tree for the Romanian gapping with *pro*-drop

All syntactic and semantic information (cf. the *SYNSEM* feature abbreviated as *ss* in the tree) about remnants in the gapped clause is registered in a *CLUSTER* head feature (proposed independently by Mouret 2006 to account for non-constituent coordinations in general). Accordingly, the sequence of the two remnants, the NP *sora mea* ‘my sister’ and the PP *la muzeu* ‘at the museum’, syntactically forms a non-headed ‘cluster’ phrase, whose category is underspecified (*XP*), thus allowing the combination with forms such as *and not* or *but not*, which are incompatible with a finite clause (as mentioned above). The combinatorial potential of a cluster is restricted, so that clusters never occur on the argumental structure of a predicate.

At a higher level, there is a head-only *fragment* phrase (cf. Ginzburg & Sag 2000), inheriting from its single head daughter (i.e., the cluster phrase) its underspecified category. A context feature (*CTXT*) contains information about the source clause and about the correlates. As far as syntactic parallelism is concerned, the fragment phrase obeys only one syntactic constraint: remnants have to unify their *HEAD* features with the *HEAD* features of their correlates in the source clause, using the contextual *SAL(ient)-UTT(utterance)* feature introduced by Ginzburg & Sag (2000). As abbreviated by the [*MAJOR+*] specification, each remnant of the target must be paired with some ‘major’ correlate in the source (namely some correlate that depends on a verbal head) and, consequently, match a possible subcategorization of the verbal predicate in the source, in accordance with Hankamer’s (1971, 1973) Major Constituent Condition.⁶ Otherwise, they may differ from their correlates with respect to their category, position or surface realization, as long as the underspecified result of the unification of their *HEAD* features matches the subcategorization requirements of the verbal predicate in the source.

Crucially, correlates are not necessarily instantiated in the syntax: they can be typed as ‘non-canonical’ and therefore realized as null pronouns, cf. *pro* for the subject correlate.⁷ This accounts for cases of gapping with *pro*-drop. The analysis of *pro*-drop in HPSG builds on the dissociation between argument structure (*ARG-ST*) and valence (Manning & Sag 1999; Ginzburg & Sag 2000):⁸ a non-canonical (*pro*-drop) argument (in our case, the subject, defined as the least oblique argument, which corresponds to the first element on the *ARG-ST* list) does not appear on a valence list, whereas a canonical realized subject is shared between *ARG-ST*

6. According to Hankamer (1973 fn. 2), “a major constituent of a given sentence S_0 is a constituent either immediately dominated by S_0 or immediately dominated by *VP*, which is immediately dominated by S_0 ”.

7. Non-canonical elements can also be realized as verbal affixes.

8. The argument structure is an ordered list of a head’s morphosyntactic dependents (=semantic arguments), and the valence is an ordered list of subject/specifiers and complements a head subcategorizes for.

and valence lists. This is illustrated in Figures 5 and 6 below, where we have two different lexical entries for the verb *merg* “(I) go” with and without an overt subject. In the canonical case of overt subject, the subject occurs both as valent and as first argument on the ARG-ST list (see the index sharing [2] in Figure 5), whereas in the case of *pro*-drop, the first argument on the ARG-ST list – a non-canonical ‘synsem’ of type *pro* – has no correspondent on the valence list (the subject list is empty in Figure 6). However, in both cases, we assume an attribute MORPH for the morphological features relevant to inflection and we identify the verb’s morphological AGR feature value with the index of the first argument on the ARG-ST list (see the index sharing [1] in both figures), in order to deal with subject-verb agreement. Overall, as previously mentioned, in this constraint-based framework, the *pro*-drop subject is not represented as a node in the tree structure (and deleted at PF), but in the lexical argument structure of the verb.

$$\left[\begin{array}{l} \text{PHON} \quad \langle \textit{merg} \rangle \\ \text{MORPH} \quad [\text{ARG } \underline{1}] \\ \text{SUBJ} \quad \langle \underline{2} \rangle \\ \text{COMPS} \quad \langle \underline{3} \rangle \\ \text{ARG-ST} \quad \langle \underline{2} \text{NP}_{\square}, \underline{3} \text{PP} \rangle \end{array} \right]$$

Figure 5. Lexical entry for the verb *merg* “I go” with overt subject

$$\left[\begin{array}{l} \text{PHON} \quad \langle \textit{merg} \rangle \\ \text{MORPH} \quad [\text{ARG } \underline{1}] \\ \text{SUBJ} \quad \langle \rangle \\ \text{COMPS} \quad \langle \underline{2} \rangle \\ \text{ARG-ST} \quad \langle \text{NP}_{\square} \textit{pro}, \underline{2} \text{PP} \rangle \end{array} \right]$$

Figure 6. Lexical entry for the verb *merg* “I go” with *pro*-drop

An additional observation about *pro*-drop concerns the distinction between finite and non-finite clauses. Here, we assume that gapping with *pro*-drop can appear only in finite clauses. In non-finite clauses, it seems that the real absence of an overt subject in the first conjunct makes the sentence severely degraded, if not ungrammatical.⁹ This is illustrated for Romanian in (12) below: if the *pro*-drop of the first conjunct subject is acceptable in the gapping construction in (12a) with a finite verb, the absence of the subject is degraded in (12b) where the coordination has as predicative verb a gerund, and even ungrammatical in (12c) where the coordination

9. We thank a reviewer who pointed this out to us. However, it is not clear if the Examples (12b–c) are indeed instances of gapping (involving verb ellipsis) or rather instances of Argument Cluster Coordination (i.e., a particular type of symmetric coordination of two clusters, in the scope of a shared predicate).

has an infinitive verb. The generalization one can make for Romanian gapping with *pro*-drop is the following: *pro*-drop is allowed in gapping constructions only when the antecedent verb carries inflection. The HPSG's approach to agreement provides a good way of modeling this observation (see the verb's morphological AGR attribute above, dealing with subject-verb agreement).

- (12) a. (Eu) Merg la film, iar Ion la muzeu.
 I go.PRS.1SG to film and Ion to museum
 "(I) go to the cinema, and Ion to the museum."
- b. Mergând ^{??}(eu) la film, iar Ion la muzeu, am fost urmăriți
 go.GER I to film and Ion to museum AUX.1 been followed.PL
 de cineva.
 by somebody
 "While I was going to the cinema and Ion to the museum, we have been followed by somebody."
- c. Ar fi de dorit de a merge *(eu) la film, iar Ion
 AUX.COND.3 be of wishing of to go I to film and Ion
 la muzeu.
 to museum
 "It would be preferable for me to go to the cinema and Ion to the museum."

Such a non-structural account of gapping constructions does not involve syntactic reconstruction of the gap, but is based on semantic reconstruction: the fragment's content (CONT in Figure 4) is built from the meaning of the source, the remnants and their correlates by some relation R_{sem} . Consequently, the fragment phrase gets a propositional content.

The overall construction (i.e., the gapping construction) is a particular type of asymmetric coordination with the main conjunct being non-elliptical and verbal, and the gapped one fragmentary and non-verbal. As illustrated in Figure 4, this *coord-ph* inherits the verbal HEAD feature of the source. While syntactic parallelism is not strict, discourse parallelism is clearly required (cf. Levin & Prince 1986; Kehler 2002): some symmetric discourse relation must hold between conjuncts, as registered in the BACKGROUND (BCKGRND) contextual feature. This accounts for the fact that gapping is felicitous with symmetric relations such as parallelism or contrast (cf. the high frequency of the contrastive conjunction *iar* in Romanian) and it is excluded with cause-effect relations (such as result, concession or condition).

5. Implications for the information structure and semantics of gapping

5.1 Implications for information structure

The experimental evidence we presented in Section 3 showing the acceptability of gapping with *pro*-drop in Romanian and Spanish challenges the assumptions made in previous research with respect to the information structure of gapping constructions.

As Grosu (1987: 451) mentions, from the earlier literature on gapping (Kuno 1976; Sag 1976), it was generally assumed that remnants and correlates in a gapping construction are focus elements, cf. Kuno's Functional Sentence Perspective Principle of Gapping (Kuno 1976: 310): "Constituents deleted by Gapping must be contextually known. On the other hand, the two constituents left behind by Gapping necessarily represent new information and, therefore, must be paired with constituents in the first conjunct that represent new information". According to Rooth (1996), focus plays a role in the grammar of ellipsis, facilitating the resolution of ellipsis. Therefore, it is not surprising that many scholars attribute a focus status to any remnant and correlate (see also Hartmann 2000 and Kuno's 'Novelty Condition', postulated by Johnson 2014).

However, Bilbiie (2017) provides empirical evidence from Romanian against a double foci analysis and in favour of a contrastive topic + focus pattern, lending support to the assumption made by Winkler (2005, 2016): "In gapping, the first remnant is a contrastive topic, the second remnant a contrastive focus" (Winkler 2016: 374). Romanian gapping is commonly used with a specific contrastive conjunction *iar* "and", which is very sensitive to information structure. Bilbiie & Winterstein (2011) and Bilbiie (2017) give three empirical arguments showing that the first remnant following the conjunction *iar* in Romanian has to be a contrastive topic, not a focus: the first remnant after *iar* (i) cannot bear prosodic stress, (ii) is generally a definite expression, and (iii) cannot be modified by focus-sensitive adverbs, which are associated with informational focus elements.¹⁰ If one follows the assumptions made by Rooth (1992, 1996) and Schwabe (2000), i.e., it is the target clause which determines the information structure of the source, we arrive at the same conclusion, namely the correlate of the first remnant after *iar* in Romanian is a topic too. All these facts support the information structure parallelism required in gapping: a remnant must have the same information status as its correlate.

Furthermore, in recent work on preferences between null and overt subjects in ellipsis resolution in Spanish, Biezma (2014) shows, based on two experimental

10. By contrast, the other (non-initial) remnant(s) can bear prosodic stress, can be modified by focus-sensitive adverbs and are not semantically constrained (they can be definite or indefinite expressions).

studies, that, in a *pro*-drop language like Spanish, the presence or absence of overt subjects provides an additional mechanism (along with intonation and word order) to indicate information status. Therefore, the preference for *pro*-drop in Spanish is related to an information structure explanation: by omitting overt subjects, speakers manifest preference for not spelling out given information (see also Alonso-Ovalle et al. 2002).¹¹

Bringing together Biezma's proposal for Spanish, empirical evidence from Romanian gapping with *iar* and our experimental findings on gapping and *pro*-drop, we have evidence for a topic-focus pattern (cf. Winkler 2005, 2016; Repp 2009; Konietzko & Winkler 2010) rather than double foci analysis for gapping in general and for gapping with *pro*-drop in particular, since the *pro*-drop subject in the source clause is necessarily a topic.

5.2 Implications for the semantic contrast

The acceptability of gapping with *pro*-drop also challenges the assumptions made with respect to the contrast notion and the alternation null/overt subjects. According to Mayol (2010), contrast is recognized in the literature as an important factor triggering the presence of overt subjects: "Overt subject pronouns ... in Romance null-subject languages become obligatory when they convey contrast, while null subject pronouns ... are generally prohibited in these contexts" (Mayol 2010: 2497).

It is commonly assumed (Sag 1976; Hartmann 2000; Repp 2000, among others) that, in gapping, each remnant must stand in semantic contrast with respect to a correlate in the source; consequently, a gapping construction has to have (at least) two contrastive pairs. An appropriate contrast can only be established between elements of a well-defined alternative set (different individuals, objects, times, locations, etc.).¹² Again, the empirical evidence for this double-contrastiveness constraint comes from the high frequency of the conjunction *iar* in Romanian

11. Biezma (2014) goes further and assumes that, on the other hand, the presence of overt subjects is perceived as informationally marked, i.e., overt subjects are focused in a language like Spanish (cf. the *pro*-drop hypothesis, Biezma 2014: 92, 123). However, as far as gapping is concerned, we do not have solid evidence to consider any overt (non-*pro*-drop) subjects as foci. In addition to Romanian data with the conjunction *iar*, one of the reviewers notices that we have evidence from languages in which the subject can be dislocated, suggesting that even overt subjects are contrastive topics, e.g., French in (i) below. This is because dislocations are generally assumed to be topics.

(i) *Marie, elle parle anglais et Jeanne allemand.*
 "Marie, she speaks English, and Jeanne German."

12. See details in Bilbiie (2017).

gapping, which requires, regardless of gapping, the presence of two contrastive pairs (Bilbúie & Winterstein 2011). However, our data on gapping with *pro*-drop show that the contrast can sometimes be established implicitly, rather than explicitly: in a *pro*-drop structure, there is an implicit correlate, which is not phonetically realized. Therefore, we have to assume that contrastive pairs may have implicit elements.

Finally, the assumption made in the literature with respect to the correlation between contrast and overt elements (cf. Mayol 2010) turns out not to be empirically adequate.¹³ In cases of gapping with *pro*-drop, contrastive subjects are not obligatorily overt (as shown by our experimental results); on the contrary, *pro*-drop subjects are not prohibited in such ‘contrastive’ structures.

6. Conclusions

In this contribution, we discussed parallelism constraints in gapping and we provided empirical evidence from two acceptability judgment tasks in Romanian and Spanish, that syntactic parallelism in gapping is less strict than traditionally assumed. Parallelism constraints are stronger at the discourse and information structure levels than at the syntactic level. The only syntactic constraint stipulates that each remnant matches a possible subcategorization of the verbal predicate its correlate depends on.

Our experimental findings show that gapping with *pro*-drop is as acceptable as without *pro*-drop, challenging the double foci analysis and supporting the topic-focus pattern for gapping constructions. On the semantic level, our data show that the contrast in gapping is not always established explicitly; more generally, our data show that one cannot make a correlation between contrast and subject (null vs overt) expression in Romance. All these empirical facts are compatible with a constructional fragment-based analysis, with semantic reconstruction of ellipsis.

13. As noted by a reviewer, the same kind of argument can be made not only in relation with null vs overt subjects, but also in relation with the distinction between weak and strong pronouns. In Romance languages, we can find cases where a syntactically realized remnant has as correlate a weak pronoun, with an affixal status. See, e.g., the attested Spanish example in (i) from the CORLEC spoken Spanish Corpus, where the remnant *a él* has as correlate in the source the weak pronoun *la*. For more examples and discussion on weak pronouns in gapping contexts, see Abeillé et al. (2014) and Bilbúie (2017).

(i) CONV 033A (García-Marchena 2018: 177)

Pero el chico la ama y dicen que ella a él.
 but the boy her love.3SG and say.3PL that she DOM he
 “But the boy loves her and they say she him.”

Acknowledgments

We are grateful to two anonymous reviewers for their very detailed and insightful comments. We wish to thank also Anne Abeillé for her availability and feedback, as well as the editors of the present volume. All errors or misunderstandings are our own.

Funding

This research was supported by the LabEx Empirical Foundations of Linguistics (EFL) project, strand 2 Experimental grammar from a cross-linguistic perspective, operation SA2 Experimental syntax (Laboratoire de Linguistique Formelle – UMR 7110 CNRS et Université Paris-Diderot).

References

- Abeillé, Anne, Gabriela Bilbiie & François Mouret. 2014. A Romance perspective on gapping constructions. In Hans Boas & Francisco González-García (eds.), *Romance perspectives on construction grammar* [Constructional approaches to language], 227–267. Amsterdam: John Benjamins. <https://doi.org/10.1075/cal.15.07abe>
- Alonso-Ovalle, Luis, Susana Fernández-Solera, Lyn Frazier & Charles Clifton. 2002. Null vs. overt pronouns and the topic-focus articulation in Spanish. *Journal of Italian Linguistics* 14(2). 151–169.
- Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278. <https://doi.org/10.1016/j.jml.2012.11.001>
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48. <https://doi.org/10.18637/jss.v067.i01>
- Biezma, María. 2014. Multiple focus strategies in *pro*-drop languages: Evidence from ellipsis in Spanish. *Syntax* 17(2). 91–131. <https://doi.org/10.1111/synt.12014>
- Bilbiie, Gabriela. 2017. *Grammaire des constructions elliptiques: Une étude comparative des phrases sans verbe en roumain et en français*. Berlin: Language Science Press.
- Bilbiie, Gabriela & Grégoire Winterstein. 2011. Expressing contrast in Romanian: The conjunction *iar*. In Janine Berns, Haïke Jacobs & Tobias Scheer (eds.), *Romance languages and linguistic theory 2009: Selected papers from 'Going Romance' Nice 2009*, 1–18. Amsterdam: John Benjamins. <https://doi.org/10.1075/rllt.3.01bil>
- Bilbiie, Gabriela & Israel de la Fuente. 2019. Can gapping be embedded? Experimental evidence from Spanish. *Glossa: A Journal of General Linguistics* 4(1). 1–39. <https://doi.org/10.5334/gjgl.782>
- Bilbiie, Gabriela, Israel de la Fuente & Anne Abeillé. 2019. Embedded gapping? A cross-linguistic perspective. Poster presented at the 32nd Annual CUNY Conference on Human Sentence Processing, Boulder, CO, March 29–31.
- Bresnan, Joan. 2001. *Lexical-functional syntax* (1st edn.). Oxford: Wiley-Blackwell.

- Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2015. *Lexical-functional syntax* (2nd edn.). Oxford: Wiley-Blackwell. <https://doi.org/10.1002/9781119105664>
- Carlson, Katy. 2001. The effects of parallelism and prosody in the processing of gapping structures. *Language and Speech* 44(1). 1–26. <https://doi.org/10.1177/00238309010440010101>
- Carlson, Katy. 2002. *Parallelism and prosody in the processing of ellipsis sentences*. New York: Routledge.
- Centeno, Naiara. 2011. Gapping and determiner sharing in Spanish. Doctoral dissertation, University of the Basque Country.
- Chomsky, Noam. 1957. *Syntactic structures*. The Hague: Mouton. <https://doi.org/10.1515/9783112316009>
- Culicover, Peter & Ray Jackendoff. 2005. *Simpler syntax*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199271092.001.0001>
- Dagnac, Anne. 2016. Gapping also needs vP-coordination: An argument from French NPI licensing. *The Linguistic Review* 33(4). 503–530. <https://doi.org/10.1515/tlr-2016-0014>
- Drummond, Alex. 2013. Ixex farm. <http://spellout.net/ibexfarm/> (20 August, 2018).
- Filiaci, Francesca, Antonella Sorace & Manuel Carreiras. 2014. Anaphoric biases of null and overt subjects in Italian and Spanish: A cross-linguistic comparison. *Language and Cognitive Processes* 29(7). 825–843.
- Frazier, Lyn, Lori Taft, Tom Roeper, Charles Clifton & Kate Ehrlich. 1984. Parallel structure: A source of facilitation in sentence comprehension. *Memory & Cognition* 12. 421–430. <https://doi.org/10.3758/BF03198303>
- Frazier, Lyn, Alan Munn & Charles Clifton. 2000. Processing coordinate structures. *Journal of Psycholinguistic Research* 29(4). 343–370. <https://doi.org/10.1023/A:1005156427600>
- García-Marchena, Oscar. 2018. Polar verbless clauses and gapping subordination in Spanish. In Eric Fuß, Marek Konopka, Beata Trawiński & Ulrich H. Wäßner (eds.), *Grammar and corpora 2016*, 169–182. Heidelberg: Heidelberg University Publishing.
- Ginzburg, Jonathan & Ivan A. Sag. 2000. *Interrogative investigations: The form, meaning and use of English interrogatives*. Stanford: CSLI Publications.
- Grosu, Alexander. 1985. Subcategorization and parallelism. *Theoretical Linguistics* 12(1). 231–240. <https://doi.org/10.1515/thli.1985.12.51.231>
- Grosu, Alexander. 1987. In René Dirven & Vilém Fried (eds.), On acceptable violations of parallelism constraints. *Functionalism in linguistics*, 425–457. Amsterdam: John Benjamins. <https://doi.org/10.1075/llsee.20.23gro>
- Hankamer, Jorge. 1971. Constraints on deletion in syntax. Doctoral dissertation, Yale University.
- Hankamer, Jorge. 1973. Unacceptable ambiguity. *Linguistic Inquiry* 4(1). 17–68.
- Hartmann, Katharina. 2000. *Right node raising and gapping. Interface conditions on prosodic deletion*. Amsterdam: John Benjamins. <https://doi.org/10.1075/z.106>
- Jackendoff, Ray. 1977. *X' Syntax*. Cambridge, MA: MIT Press.
- Johnson, Kyle. 2009. Gapping is not (VP)-ellipsis. *Linguistic Inquiry* 40. 289–328. <https://doi.org/10.1162/ling.2009.40.2.289>
- Johnson, Kyle. 2004. *In search of the English middle field*. University of Massachusetts Amherst. 2004.
- Johnson, Kyle. 2014. *Gapping*. University of Massachusetts Amherst. 2014.
- Kehler, Andrew. 2002. *Coherence, reference and the theory of grammar*. Stanford: CSLI Publications.
- Konietzko, Andreas & Susanne Winkler. 2010. Contrastive ellipsis: Mapping between syntax and information structure. *Lingua* 120. 1436–1457. <https://doi.org/10.1016/j.lingua.2008.08.009>

- Kuno, Susumu. 1976. Gapping: A functional analysis. *Linguistic Inquiry* 7(2), 300–318.
- Levin, Nancy & Ellen Prince. 1986. Gapping and causal implicature. *Papers in Linguistics* 19, 351–364. <https://doi.org/10.1080/08351818609389263>
- Manning, Christopher & Ivan A. Sag. 1999. Dissociations between argument structure and grammatical relations. In Gert Webelhuth, Jean-Pierre Koenig & Andreas Kathol (eds.), *Lexical and constructional aspects of linguistic explanation*, 63–77. Stanford: CSLI Publications.
- Mayol, Laia. 2010. Contrastive pronouns in null-subject Romance languages. *Lingua* 120(10), 2497–2514. <https://doi.org/10.1016/j.lingua.2010.04.009>
- Merchant, Jason. 2001. *The syntax of silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Merchant, Jason. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27(6), 661–738. <https://doi.org/10.1007/s10988-005-7378-3>
- Mouret, François. 2006. A phrase structure approach to argument cluster coordination. In Stefan Müller (ed.), *Proceedings of the Conference on Head-Driven Phrase Structure Grammar (HPSG'06)*, 247–267. Stanford: CSLI Publications.
- Pullum, Geoffrey K. & Arnold M. Zwicky. 1986. Phonological resolution of syntactic feature conflict. *Language* 72, 751–773. <https://doi.org/10.2307/415171>
- R Development Core Team. 2008. *R: A language and environment for statistical computing*. Vienna: R foundation for statistical computing. <http://www.R-project.org/>
- Repp, Sophie. 2009. *Negation in gapping*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199543601.001.0001>
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris. <https://doi.org/10.1515/9783110883718>
- Rizzi, Luigi. 1986. Null objects in Italian and the theory of pro. *Linguistic Inquiry* 17(3), 501–557.
- Rooth, Mats. 1992. A theory of focus interpretation. *Natural Language Semantics* 1, 75–116. <https://doi.org/10.1007/BF02342617>
- Rooth, Mats. 1996. Focus. In Shalom Lappin (ed.), *Handbook of contemporary semantic theory*, 271–297. Oxford: Blackwell.
- Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT.
- Ross, John Robert. 1970. Gapping and the order of constituents. In Manfred Bierwisch & Karl Erich Heidolph (eds.), *Progress in linguistics*, 249–259. The Hague: Mouton. <https://doi.org/10.1515/9783111350219.249>
- Sag, Ivan A. 1976. Deletion and logical form. Doctoral dissertation, MIT.
- Sag, Ivan A., Gerald Gazdar, Thomas Wasow & Steven Weisler. 1985. Coordination and how to distinguish categories. *Natural Language & Linguistic Theory* 3, 117–171. <https://doi.org/10.1007/BF00133839>
- Schachter, Paul. 1977. Constraints on coordination. *Language* 53, 86–103. <https://doi.org/10.2307/413057>
- Schwabe, Kerstin. 2000. Coordinate ellipsis and information structure. In Kerstin Schwabe & Ning Zhang (eds.), *Ellipsis in conjunction*, 247–269. Tübingen: Niemeyer. <https://doi.org/10.1515/9783110952155-013>
- Silva, Gonçalo. 2014. Gapping in European Portuguese. In Martin Kohlberger, Kate Bellamy & Eleanor Dutton (eds.), *Proceedings of the 22nd Conference of the Student Organization of Linguistics in Europe (ConSOLE XXII)*, 156–166. Leiden: The Leiden University Center for Linguistics.
- Steedman, Mark. 1990. Gapping as constituent coordination. *Linguistics and Philosophy* 13(2), 207–263. <https://doi.org/10.1007/BF00630734>

- Steedman, Mark. 2000. *The syntactic process*. Cambridge, MA: MIT Press.
- Tomioka, Satoshi. 2003. The semantics of Japanese null pronouns and its cross-linguistic implications. In Kerstin Schwabe & Susanne Winkler (eds.), *The interfaces: Deriving and interpreting omitted structures*, 321–339. Amsterdam: John Benjamins.
<https://doi.org/10.1075/la.61.16tom>
- Williams, Edwin. 1978. Across the board rule application. *Linguistic Inquiry* 9. 31–43.
- Winkler, Susanne. 2005. *Ellipsis and focus in generative grammar*. Berlin: Mouton de Gruyter.
<https://doi.org/10.1515/9783110890426>
- Winkler, Susanne. 2016. Ellipsis and information structure. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 359–382. Oxford: Oxford University Press.

On focal and *wh*-projections, indirect *wh*-questions, and quantificational chains

Caterina Bonan

University of Cambridge

In this chapter, I investigate the cartography of focus using novel data from non-standard Italian (non-StandardIT) and Trevigiano, a Venetan dialect. I argue that focus is less constrained in indirect *wh*-questions in these varieties than in Standard Italian (StandIT). Indeed, in Trevigiano both focalised objects and adverbials are felicitous in constructions with a lower *wh*-phrase. Using Featural Relativized Minimality, I argue that in the case of direct objects, the problem of crossing chains is circumvented using an IP-internal clitic that absorbs the [+N] feature of the direct object. Then, I explain the behaviour of focalised adverbials in these varieties claiming that they are externally-merged directly in the Left Periphery, hence create a high focus-chain that does not interfere with the creation of the lower *wh*-chain.

Keywords: syntax, cartography, focus, peripheries, *wh*-phrases, *wh*-movement, clitic doubling, chain formation, Romance

1. Focal projections in the cartographic enterprise

1.1 The left periphery of the clause

Since Rizzi (1997), the Left Periphery (LP) of the clause has been taken to consist of strictly-ordered projections that encode functional information such as force, finiteness etc. These projections are delimited by ForceP, where a connection is established between the clause and the discourse or a higher selecting V, and FinP, which is in direct contact with the IP, as shown in (1):

$$(1) \text{ [}_{\text{ForceP}} \text{ [Force}^\circ \text{ [}_{\text{TopP}^*} \text{ [Top}^\circ \text{ [}_{\text{FocP}} \text{ [Focus}^\circ \text{ [}_{\text{TopP}^*} \text{ [Top}^\circ \text{ [}_{\text{FinP}} \text{ [Fin}^\circ \text{ [}_{\text{IP}} \text{ [I}^\circ \text{ ...]]]]]]]]]]]]] \quad (\text{Rizzi 1997: 297 Example 41})$$

of focus' is preserved, and no interpretive clash arises (see Bianchi et al. 2017 for a finer-grained semantic explanation).

Syntactically, in indirect focus-containing *wh*-questions the processes of chain formation are complex, since in most cases the focus-chain and the *wh*-chain cross over each other (Rizzi 2001). In this chapter, I argue that languages such as non-StandIT and Trevisiano, a Venetan dialect, circumvent all chain-formation problems by means of special syntactic properties such as the availability of clitic doubling for foci, and the presence of left-peripheral adverbs whose focalisation does not disturb the formation of the *wh*-chain.²

1.2 The low periphery of the clause

Along with the LP of the clause, some authors argue that there exists a low periphery, that of *vP/VP* (Belletti 2004). This clause-internal periphery contains (minimally) the projections in (5), all of which have a discourse-related nature:

- (5) ... [_{TopP} Top° [_{Foc} Foc° [_{TopP} Top° ... VP]]] (Belletti 2005: 9)

The periphery of *vP/VP* exhibits a design that closely resembles the clause-external LP. For example, from the beginning of the cartographic enterprise, it has been clear that languages like Italian have lower 'in situ' focalization (Cinque 1993; Rizzi 1997), as suggested by examples such as (6):

- (6) *Ho letto IL TUO LIBRO (, non il suo).*
 "I read YOUR BOOK, not his." (Rizzi 1997: 287 Example (7))

Starting from Belletti (2004), instances of low focus such as the one in (6) have been argued to target the *vP/VP*-peripheral Foc. However, while both *wh*-phrases and foci compete for the Spec position of the left-peripheral FocP (Rizzi 1997), the *vP/VP*-peripheral focal projection appears only available for foci, at least in Standard Italian (StandIT). Indeed, *wh*-in situ is a question-formation strategy that StandIT rejects. In § 2, I claim that evidence from Trevisiano and non-StandIT suggest that the *vP/VP*-peripheral focal projection is actually able to host *wh*-phrases.

2. The variety of Trevisiano described here is the one spoken in a 'mixed' area that has both Sinistra Piave ('left bank' of the Piave river) and Destra Piave ('right bank') features. This variety is described and discussed in detail in Bonan (2019). As for non-standard Italian, throughout this chapter this term is used to refer mainly to the regional variety of Italian spoken in the Veneto region, unless otherwise specified.

2. Foci and *wh*-phrases

In Trevigiano, foci and fronted *wh*-phrases display the properties and distribution of their StandIT counterparts, as described in Cinque (1990) and Rizzi (1997), but not exclusively. Indeed, the focal elements of Trevigiano (and non-StandIT) display two interesting distributional properties that are not observed in StandIT. In what follows, I shall only discuss what makes Trevigiano peculiar with respect to StandIT.

The first difference concerns the availability of Belletti's (2004) clause-internal Foc for *wh*-phrases. As argued at length in Bonan (2019), the clause-internal *wh*-items of Trevigiano systematically target a linear position right below the finite V in T° (or the past participle, which in turn moves out of vP, in an Italian-like fashion). Since Trevigiano is an SVO language in which adverbials must always follow all arguments selected by the verb, this clause-internal movement of *wh*-items is only visible in the cases of *wh*-IOs and *wh*-adverbs, as illustrated in (7).³ In the absence of clause-internal movement in interrogatives, i.e., if the unmarked declarative order is conserved, ungrammaticality is observed, as in (8):

- (7) a. *Ghe ga-tu dato a chi_i a tecia _____i?*
 DAT have=you given to who the saucepan
 "Who did you give the saucepan to?"
 b. *Ga-tu magnà cuando_i el dolse _____i?*
 have=you eaten when the cake
 "When did you eat the cake?"
- (8) a. **Ghe ga-tu dato a tecia a chi?*
 DAT have=you given the saucepan to who
 b. **Ga-tu magnà el dolse cuando?*
 have=you eaten the cake when

I argue that the clause-internal *wh*-items of Trevigiano undergo short, IP-internal movement.⁴ Following Kato's (2013) work on Brazilian Portuguese and an intuition briefly sketched in Manzini (2014), I show that the targeted projection is Belletti's (2004) Foc. As a consequence, the derivation of sentences like (7) is roughly done along the lines of (9):

3. Note that the DOs in the examples in (7) are *not* right dislocated – in Trevigiano right dislocation can only exist in constructions with clitic doubling and in the presence of so-called 'comma intonation' (that is, when the dislocated chunk is phrased as an independent intonational phrase). Also, Italian-like 'marginalization' (Antinucci & Cinque 1977; Cardinaletti 2002; Samek-Lodovici 2015, among others) is categorically excluded in this variety (Bonan 2018, 2019).

4. But see Bonan (2019) for a claim that the movement under investigation is actually an instance of focus movement à la Kahnemuyipour (2001) (or of 'non-*wh*-movement' in Bošković's 1997 words).

- (9) ... [_{Foc} *wh*-item_i [*wh*^o [TopicP [Top^o [_{vP/VP} [V^o _____i]]]]]]

The Trevisano data suggest that both the clause-external and the clause-internal focal projections can indeed host not only foci but also *wh*-words. It follows that the impossibility for StandIT to license *wh*-in situ must depend on language-internal properties, not on an inherent incapacity of the *vP/VP*-peripheral Foc to host *wh*-words. The second major divide between Trevisano (and non-StandIT) and StandIT concerns the productivity of focus inside indirect *wh*-questions, which is the subject of § 2.1.

2.1 On Foc>Wh

Since Rizzi (1997), it has been widely acknowledged that the co-occurrence of focus and a *wh*-phrase is ruled out from matrix questions, in any order. This is illustrated by the StandIT examples in (10):

- (10) a. **A chi IL PREMIO NOBEL dovrebbero dare?*
 “To whom THE NOBEL PRIZE should they give?”
 b. **IL PREMIO NOBEL a chi dovrebbero dare?*
 “THE NOBEL PRIZE to whom should they give?”
 (Rizzi 1997: 298 Exmpl 45)

In contrast, as already seen in (3a), a focus and a *wh*-item can marginally co-occur in indirect *wh*-questions *iff* the focalised element is an IO. Indeed, the combination of contrastive foci and *wh*-items of any other nature is ungrammatical, as illustrated by Rizzi’s example in (11):

- (11) **Mi domando QUESTO a chi abbiamo detto.*
 “I wonder THIS to whom they have said.”
 (adapted from Rizzi 2001: 4 Example (14))

2.1.1 Trevisano and non-standard Italian

In contrast to what is observed in StandIT, a focalised DO is felicitous in constructions with a lower *wh*-item in Trevisano, as long as the DO is construed with a co-indexed clause-internal clitic. I illustrate this in (12):⁵

5. *Ghe* (“to her/him”, dative marker) might seem to be a doubling element. I argue it is not, since in the variety of Trevisano described here its presence is always compulsory in constructions with datives (it is a classic instance of Romance clitic doubling, as described for French and Spanish in Kayne 1991 and Uriagereka 1996, respectively). In the closely-related varieties that do not display clitic doubling of datives, all of the examples in this section would systematically lack *ghe*.

- (12) a. *Vojo saver EL GATO_i a chi che i gheo_i gà dato.*⁶
 want know THE CAT to who that they DAT.it gave
 “I want to know THE CAT to whom they gave (it) (not the dog).”
- b. *Me domando TO MARE_i dove che i a_i gà vista.*
 I wonder YOUR MOM where that they her have seen.F
 “I wonder YOUR MOM where they saw (her) (not your father).”

Let us call the constructions in (12) Clitic-Doubled Foci (Cl-DF). In the absence of the clause-internal clitic, ungrammaticality arises, as in (13):

- (13) a. **Vojo saver EL GATO a chi che i ghe gà dato.*
 want know THE CAT to who that they DAT gave
- b. **Me domando TO MARE dove che i gà visto.*
 I wonder YOUR MOM where that they have seen

Cl-DF are also available in non-standard Italian, in constructions that, again, crash in the absence of clitic doubling. An example is provided in (14):

- (14) *Mi domando TUA MADRE dove *(l)'hanno vista.*
 I wonder YOUR MOTHER where her=have_{3PP} seen
 “I wonder YOUR MOTHER where they saw (her) (not your father).”

In Trevigiano, focalised IOs construed with a lower *wh*-phrase are the only arguments that do *not* require clitic doubling (other than the compulsory *ghe*), as in (15):

- (15) *Me domando A TO MAMA chi che ghe gà dato da bevar.*
 I wonder TO YOUR MOM who that DAT gave to drink
 “I wonder TO YOUR MOM who gave drinks (not to your dad).”

Similarly, in the indirect *wh*-questions of Trevigiano an adverbial focus is never construed with a clitic either (16). This fact is unsurprising, since Trevigiano lacks all adverbial clitics, but needs to be accounted for:

- (16) *Me domando DOPO SENA chi che te gà visto.*
 I wonder AFTER DINNER who that you saw
 “I wonder AFTER DINNER who you saw (not this morning!)”

Again, constructions like (15) and (16) are also possible in non-standard Italian.

6. Here, and in most examples that follow, I am leaving out the (optional) negative tag for reasons of space. I shall nonetheless include one in each translation, to make sure the constructions are understood correctly.

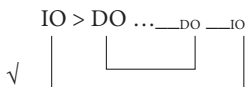
3. An analysis of clitic-doubled foci

3.1 Enter featural relativized minimality

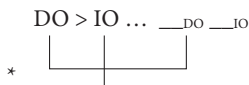
The co-occurrence within the same sentence of two quantificational elements like focus and a *wh*-phrase is theoretically problematic. Indeed, it is not clear how the focus-chain and the *wh*-chain can form properly, at least if we take both elements to start out clause-internally.

Rizzi (2001) explains the felicity, in the indirect *wh*-questions of StandIT, of focalised-IO > *wh*-DO (3a) and the ungrammaticality of focalised-DO > *wh*-IO (11) in terms of the crossing constraint à la Pesetsky (1982): the relevant A'-dependencies are nested in the former and crossed in the latter, and only nesting is unproblematic. I illustrate this in (17):

- (17) a. Nesting chains (IO > *wh*-DO):



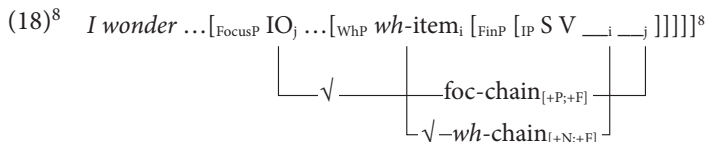
- b. Crossing chains (DO > *wh*-IO):



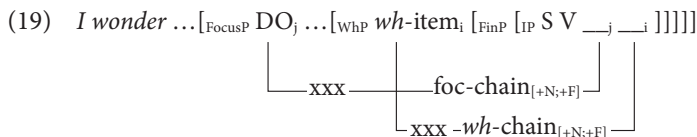
However, an analysis in terms of the crossing constraint is not sufficient to explain all data considered here. Indeed, the crossing chains appear to be felicitous in varieties such as Trevigiano and non-standard Italian, in which structures such as DO > *wh*-IO/*wh*-Adv and DO/IO > *wh*-Adv are fine.

As a consequence, I claim that the felicity of indirect *wh*-questions containing a focalised IO (and eventually of any focus-containing *wh*-question with crossing interpretational chains) is better explained if Featural Relativized Minimality (fRM) is used instead of the crossing constraint. Indeed, in experimental work, Friedmann et al. (2009), and later Villata et al. (2016), convincingly argued that most cases of intervention can be accounted for in terms of features. Accordingly, no intervention effects occur when two elements are associated to different features (a configuration known as ‘feature disjunction’), weak effects occur when the two elements share only one of two features (‘feature inclusions’), and ungrammaticality arises when the two elements carry the exact same features (‘feature identity’).

Ever since Cinque (1990), PPs have been known to have different, less constrained movement properties compared to NPs, which makes it tempting to link the special behaviour of the IOs of StandIT to the presence of (minimally) a [+P] feature, as opposed to a [+N] feature on the *wh*-item. In the framework of fRM, the [+P;+F]⁷ chain created by a focalised IO enters a relation of feature inclusion with the [+N;+F] *wh*-chain, creating minimal (to no) intervention effects when the two cross. This is sketched in (18):



On the other hand, a focalised DO is ruled out from the indirect *wh*-questions of Italian because its [+N;+F] features enter a configuration of ‘feature identity’ with the [+N;+F] features of the *wh*-phrase, hence the two chains are not able to form correctly. This is illustrated in (19):



Note that, in this theoretical framework, the ungrammaticality of examples in which a [+N;+F] focalised DO co-exists with a lower *wh*-IO is explained *iff* one takes the *wh*-phrases of Romance to be adjoined within complex *wh*-NPs, i.e., to have (silent) nominal features even in constructions with a P, along the lines of (20):



I believe that the analysis of the internal structure of Romance *wh*-PPs in (20) is legitimated by the fact that Romance languages, contrary to many Germanic languages including English, do not allow preposition stranding. An example from French is provided in (21):

7. I use ‘F(eature)’ here to refer to any feature(s) that might be responsible for the movement of *wh*-phrases and foci in interrogatives (+int, +foc, +q, etc.).

8. Note that I purposefully draw this process of chain formation, which in Rizzi’s terms is explained in terms of nesting, as an instance of chain crossing. Indeed, I do believe that any case of nesting could technically be analysed as crossing instead, as the sketch in (18) illustrates.

If my intuition is correct, the different distribution of adverbial foci in the indirect *wh*-questions of StandIT and of non-standard varieties can be linked to the different availabilities of adverbials externally-merged in the LP.

However, none of these explain the role played by the clitic in Cl-DF, namely constructions in which a focalised DO co-occurs with a lower *wh*-phrase and is construed with a co-indexed clitic. I address this in the section that follows.

3.3 Division of labour between the focalised NP and its doubling clitic

Since Cinque (1990) and Rizzi (1997), one of the major divides between the topics and foci of StandIT has been identified in the incompatibility of the latter with clitic doubling, as illustrated in (23):

- (23) *IL TUO LIBRO* (*lo) *ho comprato* (, *non il suo*).
 “YOUR BOOK I bought (not his).” (Rizzi 1997: 290 Example (16))

In contrast, topics are systematically construed with a clitic, as in (24):

- (24) *Il tuo libro*, (*lo) *ho comprato*.
 “Your book, I bought it.” (Rizzi 1997: 289–290 Example (15))

The contrast between (23) and (24) makes the case of Cl-DF theoretically challenging. Indeed, the first question that needs addressing is whether these are real foci or rather somewhat topical elements.

3.3.1 *Cl-DF are real foci*

Valentina Bianchi (p.c.) notes that indirect *wh*-questions like (11) sound significantly better if the focus is construed with a clitic, as illustrated in (25):

- (25) *Mi domando QUESTO a chi l'abbiano detto* (*non qualcosa d'altro*).
 I wonder THIS to who it'have_{SUBJ} said (not something else)
 “I wonder THIS to whom they have said (it) (not something else).”

Also, Bianchi stressed that, despite the presence of the clitic, (25) is not an instance of contrastive topicalization, as witnessed by the possible presence of a negative tag such as *non qualcosa d'altro* (“not something else”). In contrast, a negative tag is never legitimately construed with a contrastive topic, as in (26):

- (26) *Questo, l'ho già detto a Gianni* (**non quest'altro*).
 this it'have already told to John not something else
 “This, I have already said (it) to Gianni (*not this).”

The property of foci discussed in (25) also applies to all cases of Cl-DF discussed for Trevigiano and non-StandIT in this chapter. Indeed, the availability of foci construed with clitic doubling is a rather robust phenomenon in Romance and constitutes a solid test for focushood.⁹

Importantly, one must note that in Trevigiano and non-StandIT matrix foci are systematically incompatible with clitic doubling, as in StandIT (Cinque 1990; Rizzi 1997). Indeed, Trevisian non-doubled foci display all classic properties that set Italian matrix foci apart from topics, as discussed in Rizzi (1997): (i) incompatibility with clitic doubling, (ii) weak cross-over effects, (iii) incompatibility with *wh*-phrases, and (iv) resistance to reiteration (Bonan 2019). It follows that the clause-internal clitic of Cl-DF is a strategy that is peculiar to the specific syntactic context of indirect *wh*-questions, not one that is independently needed by contrastive foci. The clitic notwithstanding, Cl-DF also displays the classic properties of matrix foci. Therefore, despite the presence of a clitic, Cl-DF is not to be considered an instance of topicalization.

3.3.2 *On the role of the clause-internal clitic*

I have extensively claimed that, in Trevigiano and non-StandIT, the co-occurrence of a focus and a *wh*-phrase is less restricted than in StandIT in the sense of Rizzi (1997). However, that in Italian the felicity of a focalised element α crossing over another focal element β can depend on the presence of a clitic co-indexed with α had been observed in Wagner (2012). Observe (27–27’):

(27) “The exam was definitely not too easy. Many problems were such that only some students could solve them.”

(27’) [α *Solo il problema più facile*] *(lo_α) *ha risolto* [β *anche lo studente peggiore*].
(Wagner 2012: 14 Example (25))

In (27’), the clitic that resumes α is crucial for the felicity of the question: *solo* (“only”) takes scope over *anche* (“also”), excluding alternative interpretations such as # *the most difficult problem was even solved by the worst student* and # *the medium difficulty problem was even solved by the worst student*. As for clitic placement, the pattern in sentences with multiple overt focus-operators, as (27’), are parallel to those of topics (Wagner 2012).

Quite clearly, at least in the case of DOs, the presence of a doubling clitic seems to be required to save structures where an operator crosses over another operator.

9. Except Romanian, in which the phenomenon of clitic doubling is actually related to the notion of specificity and, for example, *wh*-phrases are normally clitic-doubled (Dobrovie-Sorin 1990, 1994), as pointed out by an anonymous reviewer.

Indeed, the clitic also partially guarantees an amelioration effect when a *wh*-DO crosses over a focal element introduced by *anche* (“also”), as illustrated by the contrast in (28), or another *wh*-phrase, as in (29):

- (28) a. *??Quale problema ha risolto anche lo studente peggiore?*
 which problem has solved also the student worst
 “Which problem also the worst student solved?”
 b. *Quale problema lo ha risolto anche lo studente peggiore?*
 which problem it has solved also the worst student
 “Which problem also the worst student solved it?”
- (29) a. *??Quale problema ti chiedi chi possa risolvere?*
 which problem you wonder who can_{SUBJ} solve
 “Which problem do you wonder who might solve?”
 b. *?Quale problema ti chiedi chi lo possa risolvere?*
 which problem you wonder who it can_{SUBJ} solve
 “Which problem do you wonder who might solve (it)?”

Despite sounding non-standard, (28b) and (29b) are clearly less marginal than their clitic-less counterparts in (a).

I wish to analyse the clitic of Cl-DF as a syntactic device that marks the path of an operator that takes wide scope over another operator, as for example the path of a focalised DO that crosses over a *wh*-phrase. More specifically, my claim is that the clitic and the coindexed NP must start off together, as part of the same ‘big DP’: I believe it is reasonable to assume that there is some division of labour between the clitic and the NP, in the sense that the clitic carries the [+N] feature and the NP carries the [+F] feature.¹⁰ When the clitic, in order to satisfy its nature, moves to its cliticization site in the high IP (à la Pollock 1989), it carries along the [+N] feature, while the NP moves to its left-peripheral focalisation site undisturbed because it has been stripped of its nominal feature by the clitic and is hence only endowed with a [+F] feature. Consequently, the focus-chain that is created entertains a felicitous relation of feature inclusion with the *wh*-phrase. This is illustrated in (30–30’):

- (30) *Vojo saver EL GATO a chi che i gheo già dato*
 want_{1PS} know THE CAT to whom that they DAT=it gave
 “I want to know THE CAT to whom they gave it (not the dog!)”
- (30’) Input: the NP and the clitic start out as a ‘big DP’;
 a. Step 1: the clitic absorbs the [+N] feature of the NP.
 b. Step 2: the clitic moves to a dedicated ClP in the higher IP

10. I wish to thank an anonymous reviewer for this suggestion.

4. Further thoughts on chain-formation and the peripheries

In this chapter, I have argued that, in Trevigiano and in non-StandIT, focus-containing indirect *wh*-questions are less constrained than in StandIT. I explained this phenomenon as the result of the availability, in these varieties, of focalised adverbials externally-merged directly in the LP and of a clitic that absorbs the [+N] feature on the NP that undergoes focalisation. The latter I explained as a semantic tool that makes proper focus-chain formation possible, despite the presence of a *wh*-chain: once the nominal feature is stripped from the focalised chunk into the clitic, the focus-chain only carries a [+F] feature that does not disturb the creation of the [+N;+F] *wh*-chain. From this analysis, it follows that the major divide between, on the one hand, the indirect *wh*-questions of Trevigiano and similar varieties and, on the other hand, those of StandIT, lies in the unavailability of ‘big DPs’ in the latter.

4.1 On ‘mild’ feature relations

Similarly, the possibility of fronting a focalised-IO, available also in StandIT (Rizzi 1997, 2001), is attributed here to the special nature of indirect objects: PPs have been known for decades to move more freely compared to DOs and subjects (Cinque 1990), which I linked to the fact that their [+P;+F] features entertain a relation of feature inclusion with the [+N;+F] *wh*-chain. Note that, in Friedmann et al.’s (2009) terms, feature inclusion should result in a weak violation, which is not observed here. I claim that this is easily explained if one considers that the [+F] feature checked by the focus is different from the [+F] feature in WhP checked by the *wh*-phrase (in Bonan 2019, I argue that the former is [Q_{foc}] while the latter is [Q_{int}]). As a consequence, I believe that these configurations can be treated as instances of ‘mild’ feature inclusion, which results in no violation at all. In contrast, a focalised DO that crosses over a *wh*-phrase creates an undesirable [+N;+F] configuration of (mild) feature identity, which results in ungrammaticality (as in StandIT), unless the [+N] feature can be ‘stripped’ from the focus by means of a grammatical device such as the clitic of Trevigiano and non-StandIT Cl-DF, which again results in a configuration of ‘mild’ feature inclusion ([+Q_{foc}] vs [+N;+Q_{int}]).

Note that there exist (standard and non-standard) Italian cases in which a focalised S crosses over a *wh*-phrase, such as the one in (32). These, I believe, are not problematic for the theory developed here:

- (32) *Mi domando I ROSSI chi abbiano visto.*
 I wonder THE ROSSI who have_{SUBJ} seen
 ‘I wonder who THE ROSSI saw.’

Indeed, (32) is only apparently an instance of two [+N;+F] chains in a relation of feature identity: following Rizzi (2005), the nominal feature on the subject must also carry a special [+aboutness] feature, which makes the two nominal chains slightly different, hence giving rise to a relation of either feature inclusion (if we take [+aboutness] and [+N] to be separate features) or of ‘mild’ feature identity (if we take [+N] and [+N_{about}] to exist). In either case, the felicity of (33) is correctly predicted.

4.2 Peripheral focal- and *wh*-projections

For the sake of completeness, let us see how the left peripheral and *vP*/*VP*-peripheral focal- and *wh*-projections are used in the varieties under investigation. We saw that one and only one focus can appear in the LP (focus fronting) and one and only one in the *VP*-periphery (clause-internal focus). These are likely to target, respectively, Rizzi’s (1997) left-peripheral FocP and Belletti’s (2004) clause internal Foc. The incompatibility of two foci within the same sentence has been explained, since Rizzi (1997), in terms of interpretive implausibility of focus-containing foci. The same interpretation-driven incompatibility is observed in direct questions of a focus and a *wh*-phrase. Elements of these types can co-occur neither within the same periphery, as illustrated in (33), nor in different peripheries, as in (34), irrespective of their linear order and the grammatical role:

- (33) **QUESTO a chi hanno detto (, non qualcos’altro)?*
 THIS to whom have_{3PP} said (, not something else)
 (Rizzi 2001: 4 Example (13))

- (34) **Quando hanno consegnato IL LIBRO a Leo?*
 when have_{3PP} given THE BOOK to Leo (Bocci 2013: 19)

In contrast, when combining a focus and a *wh*-phrase in indirect *wh*-questions, the co-occurrence of these elements is grammatical because indirect interrogatives are not inherently information-seeking: the movement of the *wh*-phrase into Spec,WhP is indeed to be considered as mere *wh*-movement which causes no semantic clash (Bianchi et al. 2017; Bonan 2019). In the cases, discussed in § 2, in which the focus of such constructions does not undergo fronting, this can either be argued to stay in its first-merge position (in situ focus) or to undergo short movement into the Spec of Belletti’s (2004) Foc (clause-internally moved focus, see Bonan 2019 for a detailed investigation). However, Belletti’s Foc is not only able to host contrastively-focused clause internal elements (be they alone in the structure or construed with a higher *wh*-item in Spec,WhP). In Trevigiano, this *vP*/*VP*-

VP-peripheral focal projection is available for *wh*-phrases as well (Bonan 2019). This is sketched in (35):

- (35) *Si-tu* ... [_{IP} *ndà* [_{Foc} *quando*]_j ... [_{VP} *al marcà* ___]_j]]]?
 have-you gone when to.the market
 “When did you go to the market?”

Interestingly, the impossibility for StandIT to license *wh*-items clause-internally, which has been widely discussed in the literature, ceases to exist in the non-standard varieties in contact with dialects that have ‘optional *wh*-in situ’, such as Venetan Italian. Examples are provided in (36):

- (36) a. Seeing your friend’s new cardigan. You ask:
E l’hai comprata dove, questa meraviglia?
 and it’have_{2PS} bought where this wonder
 “And where did you buy this beauty?”
 b. Sarah is happy because Marco has finally called her. You ask:
Ti ha chiamata quando, Marco?
 you has_{3P} called when Marco
 “When did Marco call you?”

Plausibly, the choice between overt scope (as in StandIT, in which overt *wh*-fronting is required) and scope construal (as in Trevigiano and other optional ‘in situ’ languages) is ruled by a component of grammar that is quite vulnerable, which can explain why dialect-influenced varieties of non-standard Italian display interrogative constructions that the standard variety rejects. My claim is that the impossibility for *vP* to host *wh*-phrases in StandIT cannot be linked to an inherent incompatibility of the clause-internal Foc and *wh*-items, but to linguistic reasons that go beyond the subject of this work.

5. Conclusions

To conclude, my claim is that it is possible to explain the different distribution of focus in indirect *wh*-questions in StandIT on the one hand and Trevigiano and non-StandIT on the other, if we take them to be linked to the unavailability in the former of (i) ‘big DPs’ whose clitic is able to ‘absorb’ the [+N] feature of the focalised NP, and (ii) adverbials eligible for focalization that can be externally-merged directly in the LP. One might wonder why in languages in which ‘big DPs’ are available, clitic doubling is categorically not available in constructions with matrix focus. However, I believe that the use of such a tool would be utterly unnecessary in combination with matrix focus, and is hence ruled out by Economy.

Acknowledgments

I am indebted to Ur Shlonsky, Giuliano Bocci, Lena Baunaz, Giuseppe Samo, Valentina Bianchi and Kenneth Baclawski Jr. for their invaluable comments and suggestions. Of course, I take full responsibility for any ideas (and mistakes) endorsed here.

Funding

This work was financially supported by the Swiss National Science Foundation (SNSF): first by Ur Shlonsky and Giuliano Bocci's Geneva-based project n° 100012_156160, and then finalized during my own mobility project n° P2GEP1_184384 at the University of Cambridge.

References

- Antinucci, Francesco & Guglielmo Cinque. 1977. Sull'ordine delle parole in italiano: L'emarginazione. *Studi di grammatica italiana* 6. 121–146.
- Belletti, Adriana. 2004. Aspects of the low IP area. In Luigi Rizzi (ed.), *The structure of CP and IP*, Volume 2, 16–51. Oxford: Oxford University Press.
- Belletti, Adriana. 2005. Extended doubling and the VP periphery. *Probus* 17(1). 1–35. <https://doi.org/10.1515/prbs.2005.17.1.1>
- Bianchi, Valentina, Giuliano Bocci & Silvio Cruschina. 2017. Two types of subject inversion in Italian *wh*-questions. *Revue roumaine de linguistique* 62. 233–252.
- Bocci, Giuliano. 2013. *The syntax-prosody interface. A cartographic perspective with evidence from Italian*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.204>
- Bonan, Caterina. 2018. On insituness and (very) low *wh*-positions. The case of Treviso. *GG@G: Generative Grammar in Geneva 11. Special Issue: Proceedings of the 1st SynCart workshop "From maps to principles"*, edited by Giuseppe Samo, Karen Martini & Giuliano Bocci, 21–41.
- Bonan, Caterina. 2019. On clause-internally moved *wh*-phrases. *Wh-to-Foc, nominative clitics, and the theory of northern Italian wh-in situ*. Doctoral dissertation, Université de Genève.
- Bosković, Željko. 1997. Superiority effects with multiple *wh*-fronting in Serbo-Croatian. *Lingua* 102(1). 1–20. [https://doi.org/10.1016/S0024-3841\(96\)00031-9](https://doi.org/10.1016/S0024-3841(96)00031-9)
- Cardinaletti, Anna. 2002. Against optional and null clitics: Right dislocation vs. marginalization. *Studia Linguistica* 58. 29–57. <https://doi.org/10.1111/1467-9582.00086>
- Cinque, Guglielmo. 1990. *Types of A-bar dependencies*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 1993. A null theory of phrase and compound stress. *Linguistic Inquiry* 24. 239–298.
- Cinque, Guglielmo & Luigi Rizzi. 2010. The cartography of syntactic structures. In Bernd Heine & Heiko Narrog (eds.), *The Oxford handbook of linguistic analysis*, 65–78. Oxford: Oxford University Press.
- De Crousaz, Isabelle & Ur Shlonsky. 2003. The distribution of a subject clitic pronoun in a Franco-Provençal dialect and the licensing of pro. *Linguistic Inquiry* 34(3). 413–442. <https://doi.org/10.1162/002438903322247542>

- Dobrovie-Sorin, Carmen. 1990. Clitic doubling, wh-movement, and quantification in Romanian. *Linguistic Inquiry* 21(3). 351–397.
- Dobrovie-Sorin, Carmen. 1994. *The syntax of Romanian*. Berlin: Mouton de Gruyter.
<https://doi.org/10.1515/9783110886597>
- Friedman, Nama, Adriana Belletti & Luigi Rizzi. 2009. Relativized relatives: Types of intervention in the acquisition of A-bar dependencies. *Lingua* 119(1). 67–88.
<https://doi.org/10.1016/j.lingua.2008.09.002>
- Kahnemuyipour, Arsalan. 2001. On wh-questions in Persian. *Canadian Journal of Linguistics / Revue Canadienne de Linguistique* 46(1/2). 41–51. <https://doi.org/10.1017/S000841310001793X>
- Kato, Mary. 2013. Deriving “wh-in-situ” through movement in Brazilian Portuguese. In Victoria Camacho-Taboada, Ángel L. Jiménez Fernández, Javier Martín-González & Mariano Reyes-Tejedor (eds), *Information structure and agreement*, 175–192. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.197.06kat>
- Kayne, Richard S. 1991. Romance clitics, verb movement, and PRO. *Linguistic Inquiry* 22(4). 647–686.
- Manzini, Maria Rita. 2014. Grammatical categories: Strong and weak pronouns in Romance. *Lingua* 150. 171–201. <https://doi.org/10.1016/j.lingua.2014.07.001>
- Pesetsky, David. 1982. Paths and categories. Doctoral dissertation, MIT.
- Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar. Handbook on 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. Leiden: Brill.
- Rizzi, Luigi. 2004. Locality and left periphery. In Adriana Belletti (ed.), *Structures and beyond. The cartography of syntactic structures*, Volume 3, 3–15. Oxford: Oxford University Press.
- Rizzi, Luigi. 2005. On some properties of subjects and topics. In Laura Brugè, Giuliana Giusti, Nicola Munaro, Walter Schweikert & Giuseppina Turano (eds.), *Proceedings of the XXX Incontro di Grammatica Generativa*, 203–224. Venice: Cafoscarina.
- Rizzi, Luigi & Giuliano Bocci. 2017. The left periphery of the clause – primarily illustrated for Italian. In Martin Everaert & Henk C. Van Riemsdijk (eds.), *The Blackwell companion to syntax*, Volume 2, 1–30. Malden, MA: Blackwell.
- Rizzi, Luigi & Ur Shlonsky. 2007. Strategies of subject extraction. In Uli Sauerland & Hans-Martin Gärtner (eds.), *Interfaces + recursion = language? Chomsky’s minimalism and the view from syntax-semantics*, 115–160. Berlin: De Gruyter Mouton.
- Samek-Lodovici, Vieri. 2015. *The interaction of focus, givenness, and prosody. A study of Italian clause structure*. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780198737926.001.0001>
- Uriagereka, Juan. 1996. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26(1). 79–123.
- Villata, Sandra, Julie Franck & Luigi Rizzi. 2016. Intervention effects and relativized minimality: New experimental evidence from graded judgments. *Lingua* 179. 76–96.
<https://doi.org/10.1016/j.lingua.2016.03.004>
- Wagner, Michael. 2012. Contrastive topics decomposed. *Semantics & Pragmatics* 5. 1–54.
<https://doi.org/10.3765/sp.5.8>

Is there a dative alternation in Romanian?

Remarks on the cross-categorial variation of datives in ditransitive constructions

Alexandra Cornilescu and Alina Tigău

University of Bucharest

Against recent claims that Romance languages lack a genuine dative alternation since they lack a genuine Prepositional Dative Construction (e.g., Pineda 2012), we bring evidence that, in Romanian, even in Recipient ditransitive constructions, datives manifest either DP or PP properties. In order to establish this result, we examine both the (internal) structure of the Romanian inflectional dative, and the prepositional dative, marked by the preposition *la* “at”/“to” and show that *both* forms require a ‘dual categorial analysis’, in order to allow licensing of their case and person features. While the default interpretation of datives in Recipient ditransitive constructions is that of DPs (whence the possibility of clitic doubling (CD)), there is a class of contexts (e.g., double datives, featuring a possessive dative and a Goal/Recipient dative), where the Recipient must be projected as a PP, since otherwise it cannot be licensed. The dual categorization of the Recipient as a DP/PP proves the existence of a genuine dative alternation in Romanian.

Keywords: dative alternation, inflectional dative, prepositional dative, re-analysis

1. Background and aim

1.1 The aim of the chapter

The present chapter aims at showing the existence of a genuine dative alternation in Romanian, in the sense that the dative constituent in ditransitive constructions behaves either as a DP or as a PP, even if its form does not change. In order to establish this result, we examine both the Romanian inflectional dative (Inf=dat), and the prepositional dative (Prep=dat), marked by the preposition *la* “at”/“to” (*la*=dat) showing that *both* forms require a ‘dual categorial analysis’, in order to be properly licensed.

The opening section gives some background on the problem, and then presents the Romanian ditransitive constructions, stressing on the properties shared with other Romance languages. The second section contains a detailed presentation of the complex internal structure of dative constituents. It is shown that the main property of datives is their sensitivity to the Animacy Hierarchy (=AH), which requires the derivational valuation of a syntactic [Person] feature. The third section turns to the examination of several ditransitive configurations where the datives must be analysed as PP and cannot be CD-ed. On the other hand, in contexts where CD of the datives is available, datives behave as DPs. If the analysis is on the right track, the conclusion follows that Romanian (and other Romance languages that show similar properties) has a genuine dative alternation.

1.2 Background, properties of Romanian ditransitives

The existence of the English Double Object Construction (DOC) in Romance languages was not initially obvious (Kayne 1984), chiefly because in these languages, the nominal dative is often uniquely encoded as a PP (e.g., French, Italian, Spanish, Catalan), unlike English which shows a contrast between PP/DP encoding. However, if the existence of dative clitics is also considered, then Romance languages do show an alternation between a clitic construction, interpreted as a form of DOC, since the clitic is a DP, and a clitic-less construction, taken to be the analogue of the Prepositional Dative Construction (PDC) (e.g., Cuervo 2003; Diaconescu & Rivero 2007).

Further research showed that this correspondence is untenable, since both the clitic/ and the cliticless construction have DOC properties. The irrelevance of the clitic for binding and scope phenomena is shown by the existence of bidirectional *c*-command between DO and IO in Romance languages, as shown for French by Harley (2002), or for Spanish and Catalan by Pineda (2012, and references therein). Therefore, from a configurational point of view, there is no parallel to be found exclusively between English DOC and Romance CD-ed constructions.

Romanian ditransitives are no exception to the Romance pattern. Each argument can bind or outscope the other one, irrespective of the dative clitic, and this obtains with *la*=datives as well as with *Inf*=datives. Thus in (1), the non-doubled dative binds into the DO, and the addition of the clitic in (2) makes no difference. The second pair of examples shows that the DO has the same privileges: it may bind an undoubled goal (3), but it may also bind a doubled goal (4).

- (1) *Angajatorii nu au dat tuturor muncitorilor_i/ la toți muncitorii*
 employers.the not have given all.DAT workers.the.DAT at all workers.the
drepturile lor_i bănești.
 rights their money-related.
 “The employers didn’t give all the workers their due money.”
- (2) *Angajatorii le-au dat muncitorilor_i/ la muncitori*
 employers they.DAT.CL-have given workers.the.DAT at workers
cecurile lor_i.
 cheques their
 “The employers gave the workers their paycheques.”
- (3) *Redacția a trimis toate manuscrisele_i autorilor lor_i/*
 editorial-office.the has sent all manuscripts.the authors.the.DAT their
la autorii lor_i.
 at authors.the their
 “The editors sent all the manuscripts to their authors.”
- (4) *Redacția le-a trimis toate manuscrisele_i*
 editorial-office.the they.DAT.CL-has sent all manuscripts.the
autorilor lor_i/la autorii lor_i.
 authors.the.DAT their/at authors.the their
 “The editors sent all the manuscripts to their authors.”

Thus, in Romanian too, configurational properties (binding and scope) are unaffected by the clitic. Since the clitic is supposed to always be possible and it indicates DP status of the dative, some researchers have taken a step forward and proposed that, in Romance, dative constituents are *always* DPs, which may cliticize or be CD-ed. The following quotation is an example: “This leads us to propose that, irrespective of dative CD, Romance ditransitives (with some kind of transfer meaning) are a reflex of DOC, and that no English DPC-like ditransitive constructions exist” (Pineda 2012: 59).

Against such claims, we present evidence that at least some Romance languages (e.g., Romanian) have an alternation between PP datives and DP datives, exhibiting a genuine dative alternation.¹

1. Our proposal is reminiscent of Anagnostopoulou (2005), who argues that the category of IOs is orthogonal to their distribution in ditransitive configurations, with a DP IO being able to occur in the Prepositional Object Construction and a PP IO being able to surface in the Double Object Construction.

2. The structure of Romanian dative phrases

2.1 Inflectional and prepositional marking

A property that singles out Romanian in Romance is the presence of ‘inflectional dative morphology’, with the result that Romanian exhibits both ‘inflectional’ and ‘prepositional marking’. The preposition is *la* ‘at/to’, which has replaced the general Romance *a* (Iorga 2013).

2.2 Sensitivity to the animacy hierarchy

A crucial property of Inf=dat and Prep=dat alike is that they are sensitive to AH (5), showing a preference for human or animate noun classes.

- (5) human > animate > (abstract) > inanimate > (abstract)

Inf=dat tend to exclude inanimates, but may accept abstract nominals. *La*=dat may accept inanimates, but definitely rules out abstract nouns, as shown in (6).

- (6) a. *Am turnat vin la musafiri/musafirilor.*
 have.I poured wine at guests guests.the.DAT
 “I poured the guests wine.”
- b. *Am dat apă la cai/ ?cailor.*
 have.I poured water at horses horses.the.DAT
 “I poured water to the horses.”
- c. *Am turnat apă la flori/ *?florilor.*
 have.I poured water at flowers flowers.the.DAT
 “I watered the flowers.”
- d. *Am supus proiectul atenției boardului/ **la*
 have.I submitted project.the attention.the.DAT board.the.GEN at
atenția boardului.
 attention.the board.the.GEN
 “He submitted the project for the board’s attention.”

Actually, to accurately specify the distribution of *la*-datives, even when the description is limited to ‘standard’ Romanian, a more detailed version of AH is required (in (7), adapted from Iorga 2013: 87), on which other contrasts than in (5) can be stated. One is the contrast between pronouns and proper names versus other nominal categories. Pronominal clitics, especially 1st and 2nd person ones (8a), cannot have *la*-dative associates. Likewise, singular proper names do not have a *la*-dative form (8a). Number contrasts hold across all nominal categories: there is strong preference for using plurals over singulars (Iorga 2013 and references therein) (8b). Finally, *la*-singulars are better, if specific or definite (8c).

- (7) Pers.Prn 1st and 2nd person > Pers.Pr 3rd person > proper names > humans > animates > (abstract) > inanimates > (abstract)
- (8) a. *I-a dăruit cartea ei numai lui/ *la el/ *la Ion.*
 she.DAT.CL-has given book.the her only he.DAT at her at Ion
 “She gave her book only to him/John.”
- b. *Au dat medicamentele ?la bolnav/ la bolnavi.*
 have.they given drugs.the at sick-man at sick-men
 “They have given the medicine to the sick.”
- c. *Dă rochia ??la o fată/ ?la o fată saracă/ fetei sărace.*
 give.IMP dress.the at a girl at a girl poor girl.the.DAT poor.
 “Give the dress to a girl/to a poor girl/to the poor girl!”

These data suggest that *la*=datives, unlike *Inf*=dat, are sensitive not only to AH, but also to the Definiteness Hierarchy (DefH). One important theoretical problem is that of incorporating ‘scalar concepts’ like AH or DefH into the discrete binary system of a minimalist grammar. Richards (2008)² proposes that these hierarchies are semantic and pragmatic in nature and should be viewed as ‘syntax-semantics interface phenomena’. According to him, nouns which are sensitive to the hierarchy should be lexically specified for a binary grammatical [Person] feature. It is this feature which triggers an interpretation of the NP in particular cases. Since [Person] is a syntactic feature, not an inherent semantic one, it must be checked in the derivation.

Conclusion

Both dative markers are sensitive to the AH, but have different cut-off points regarding marking for [+Person] as apparent in (6) and (8). Datives must check a [Person] feature during the derivation.³

2. “Our claim is that [Person] in the syntax is just animacy/definiteness at the (semantic) interface. That is, we assume that there is a single, discrete, binary property [\pm Person] whose presence *vs* absence correlates with high *vs* low interpretations (on the definiteness and animacy scales) in the semantic component” (Richards 2008: 139).

The AH and DefH hierarchies enable us to draw a number of conclusions as to certain redundancy rules of the system. Firstly, it is only [+animate] NPs that may vary for person: all inanimate NPs are 3rd person and do not require any specification in this respect. Secondly, only [+definite] nominals allow for person variation, while indefinites, irrespective of whether they are specific or not, are automatically 3rd person and as such do not need any person specification.

3. As observed by a reviewer, the different distribution of *Inf*=dat and *la*=dat remains largely unexplained in the current paper. For a more thorough presentation of the Romanian dative, see Iorga (2013).

2.3 The thematic range of datives in ditransitive constructions

In spite of its lexical locative/directional meaning and in contrast to the English *to*, *la* ‘at/to’ is used not only for Goal/Recipient datives, but also for *all of the q-roles* associated with the Inf=dat phrase: Beneficiary, Maleficiary (or Source). Ficiary roles are alternatively expressed by the suitable ‘lexical’ prepositions *pentru* ‘for’ and *de la* ‘from’, which cannot be CD-ed.

- (9) Recipient/Goal/Possessor
Bunica (le)-a dat prăjituri copiilor/
 grandmother.the they.DAT.CL.-has given cakes children.the.DAT
la copii.
 at children
 ‘Grandmother gave cakes to the children.’
- (10) Beneficiary
 a. *Bunica (le)-a copt prăjituri copiilor/*
 grandmother.the they.DAT.CL.-has baked cakes children.the.DAT
la copii.
 at children
 ‘Grandmother baked some cakes for her grandchildren.’
 b. *Bunica *(le) a copt prăjituri pentru copii.*
 grandmother has baked cakes for children
 ‘Grandmother baked some cakes for the children.’
- (11) Maleficiary (Source)
 a. *Niște vagabonzi le-au furat copiilor/ la copii*
 some tramps they.CL.DAT-have stolen children.the.DAT at children
niște mere din livadă.
 some apples from orchard
 ‘Some tramps stole the children some apples from the orchard.’
 b. *Niște vagabonzi *(le)-au furat niște mere de la copii*
 Some tramps they.CL.DAT-have stolen some apples from children
din livadă.
 from orchard
 ‘Some tramps stole the children some apples from the orchard.’

Thus, both Inf=dat and *la*=dat merely signal that their referent is assumed to be human or human-like; this is the common feature of all these θ -roles. Beyond this presupposed human(-like) property, the interpretation of the *la*-phrase or of the Inf=dat mostly *depends on the descriptive content of the verb* (its a-structure). We will simply assume that this range of θ -roles is a reflex of the same [Person] feature required by the AH and propose that datives enter the derivation with an [*u*Person] feature, valued derivationally.

2.4 The internal structure of dative vs accusative *la*⁴

In this section we examine the internal structure of *la*-datives, comparing the lexical and the functional preposition *la*. Lexical *la* ‘at/to’ is the core preposition of the location and movement frames. It assigns accusative to its complement and cannot be CD-ed. Moreover, it is quite insensitive to AH or DefH:

- (12) a. *Paul *(mi)-a venit la mine.*
 Paul (I.DAT.CL) has come at I.ACC
 ‘Paul has come to me/to my place.’
- b. *(Le)-a cumpărat ciocolata asta la copii / la cofetarie*
 (they.DAT.CL)-has bought chocolate this at children at confectionery
 ‘He bought this chocolate for the children/from the confectionery.’
- c. **I-a cumpărat ciocolata asta la cofetărie.*
 him.DAT.CL -has bought chocolate this at confectionery
 ‘He has bought this chocolate from the confectionery.’

When used in ditransitive constructions, with the whole range of transfer of possession verbs, functional *la* has sharply different properties. It shares with lexical *la* only the fact that it assigns accusative to its complement. Thematically, it covers not only the θ -role Goal (with change of location verbs), but also Recipient (Possessor), Beneficiary, Maleficiary (Source), with transfer of possession verbs (as shown in (9)–(11) above), so that it may correspond to the English Ps *to*, *for*, *from*. In examples like (9)–(11), *la* has little or no descriptive content, the θ -role being entirely dependent on the selecting verb. Since one important content feature shared by Recipient (Possessor), Beneficiary, and Maleficiary, is [Person], in the sense of the AH, we propose to say that in ditransitive constructions, P *la* is bleached of meaning, retaining only a syntactic [Person] feature, which is a restriction on the object of *la*, and which is interpreted at the interface according to the AH and DefH. Thus, in ditransitive constructions, *la*-PPs are sensitive to the AH and DefH. Moreover, precisely in contexts that show sensitivity to [Person], such as ditransitive constructions, *la*-PPs can be CD-ed. Compare (12b) and (12c). In (12b), the I(ndirect) O(bject) is animate and can optionally be CD-ed. In (12c), the *la*-PP is inanimate and cannot be CD-ed. The sentence is grammatical if the *la*-PP is a [Locative] adjunct, and the clitic refers to some unnamed [Person] IO. Since Romanian has DP clitics (not PP clitics), it follows that when *la*-PPs are or may be doubled they must be analyzed as DPs. Surprisingly, these *la*-phrases are doubled

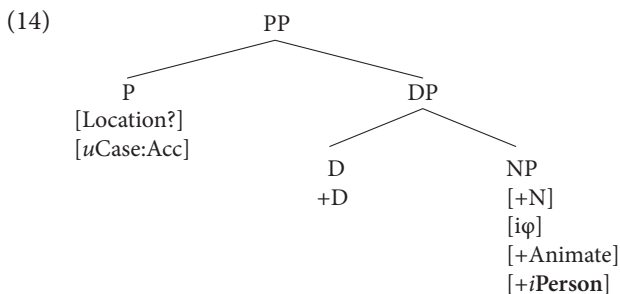
4. We thank a reviewer for calling attention to the fact that this section was so poorly written as to suggest that we consider lexical *la* and *la*-dative two distinct Ps. In fact, evidence indicates that *la*-dative is merely a (more) grammaticalized use of the same preposition, as we endeavoured to explain in the revised text.

by *dative*, not accusative clitics. Acquiring an abstract, grammatical content (i.e., [Person]), and selected by an extended class of ditransitive verbs, P *la* becomes a functional, grammaticalized category, involved in the organization of the Romanian case system. We will refer to this use of P *la* as functional *la* or *la-dative* and try to make sense of its structure.

On the strength of the properties just mentioned, a first approximation regarding the structure of the *la-dative* is (13), where the inner DP bears Acc and the outer DP bears Dat, assuming, as we must, that the clitic and its associate share the same case.

(13) [_{DP2} [_{PP} *la* [_{DP1}]]]

This representation immediately raises the question of what triggered the re-analysis of the PP as a DP. In our view, re-analysis was prompted by the need to ‘make visible’ the obligatory [Person] feature of the NP in DP₁, which had been grammaticalized, while at the same time the P had been bleached of its Location/Goal thematic content, even extending to verbs which were incompatible with Goal (e.g., *fura* ‘steal’). At this point, the interface between syntax and PF shows an imperfect correspondence between features and their exponents, since while the P had lost its θ -content, there was a syntactic feature [Person] in the DP, which was obligatory but had no PF realization (14).



This tension between syntax and PF initiates a change. *La* is re-analyzed, becoming the PF exponent of the [*i*Person] feature, easily inserted under a D head, as argued for in Longobardi (2008). As such, it may license the [*u*Person] feature that the nominal complement enters the derivation with. The D[*i*Person] head simultaneously acts as a ‘category shifter’ from PP to DP. Importantly, *la* retains its prepositional function, since it still assigns Acc case to its complement. The bigger DP₂ requires case, a case feature being projected in a higher KP, which embeds DP₂, so that the order of the projections becomes K[*u*Dative] > D[*i*Person] > P[Acc] > D[*u*Person]. What is required to signal this complex structure at PF is that the higher case assigned to DP₂ by some ‘external head’ should be different from the

lower case assigned by the Prep to DP₁. The value [Dative] of K is probably related to the [*i*Person] feature that is also part of *la*'s feature matrix.

- (15) a.
-
- ```

graph TD
 PP --> P
 PP --> DP["DP[uPerson, uAcc]"]
 P --- Case["[Case_]"]
 P --- la["la"]

```
- b.
- 
- ```

graph TD
    DP2["DP2[iPerson]"] --> D2
    DP2 --> PP
    D2 --- PD["P+D"]
    D2 --- IPerson["[iPerson]"]
    D2 --- CaseAcc["[Case:Acc]"]
    D2 --- la["la"]
    PP --> DP1
    PP --> P_prime["P'"]
    DP1 --- uPerson["[uPerson]"]
    DP1 --- Acc["[#Acc]"]
    P_prime --> P
    P_prime --> DP1_prime["DP1"]
  
```
- c.
-
- ```

graph TD
 KP["KP [uDative, +D, iPerson, + P, uAcc]"] --> K
 KP --> DP2
 K --- D2K["D2+K"]
 K --- uDative["[uDative]"]
 K --- iPerson["[iPerson]"]
 K --- CaseAcc["[Case:Acc]"]
 K --- la["la"]
 DP2 --> D2
 DP2 --> PP
 PP --> DP1
 PP --> P_prime["P'"]
 DP1 --- PersonAcc["[#Person, #Acc]"]
 P_prime --> P
 P_prime --> DP1_prime["DP1"]

```

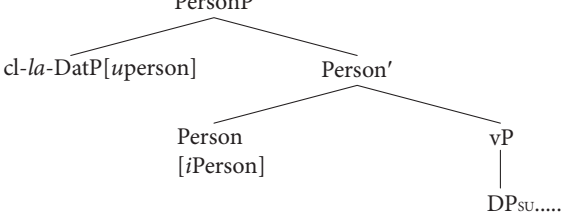
The steps of the derivation are transparent. In (15a), the lower DP<sub>1</sub> checks Acc case, assigned by [*pl*]. Next, D<sub>2</sub> [*i*Person] merges and the P *la* raises to D<sub>2</sub> becoming an exponent of Person. As shown in (15b), [<sub>D2</sub> [[<sub>P</sub> *la*]+D]] values DP<sub>1</sub>'s [*u*Person] feature], by Agree. The K head merges introducing DP<sub>2</sub>'s case feature, [*u*Dative]. The P *la* moves to K, as shown in (15c); it is now also an 'exponent', (not an assigner), of dative case. In comparison with (14), representation (15c) shows a better syntax/semantics fit, since one of the features, syncretically realized by N in (14), is realized as a separate lexical item in (15), while the second case feature signals the existence of the higher nominal category DP<sub>2</sub>/KP. Dative case is licensed by an appropriate functional head, such as an applicative head. The derivation in (15) is a hypothesis on how a P which remains an accusative assigner turns into a DP requiring the checking of a dative case feature.

*Why is the clitic possible and sometimes required?*

The presence of the clitic should be a response to some internal need of the *la*-phrase. Intuitively, and not implausibly, the former P *la*, a person/case spell-out, may be further eroded semantically, shifting from  $[la]_{\kappa} [iPerson]$  to  $[la]_{\kappa} [uPerson]$ . In this case, when there is Agree between the person head and the internal DP [*uPerson*], there is feature sharing between the [Person] head and the internal DP1, but [*uPerson*] remains unvalued, as apparent in the simplified representations below, roughly corresponding to steps (15b) and (c) from the derivation above.

- (16)  $D_2 [uPerson]$   $DP_1 [uPerson]$   $P[Case: Acc]$   $DP_1 [uPerson]$   
*la* [P+D<sub>2</sub>]

Following Belletti (2005), we assume that there is a PersonP at the  $\nu P$  periphery, targeted by CD-ed phrases on the way to T (see the motivation in Cornilescu et al. 2017). The Person [*iPerson*] head of this periphery values the [*uPerson*] of the dative. Since only CD-ed phrases reach the PersonP, the clitic becomes the spell-out of [Person], while *la*-merely spells out case.

- (17)
- 
- ```

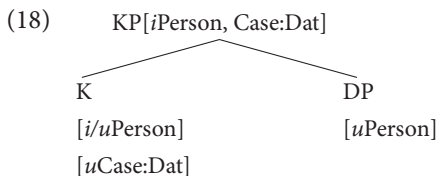
graph TD
    PersonP --> cl_la_DatP[cl-la-DatP[uPerson]]
    PersonP --> Person_prime[Person']
    Person_prime --> Person_i[Person [iPerson]]
    Person_prime --> vP[vP]
    vP --> DPsu[DPsu.....]
  
```

Conclusions

La-dat phrases are KPs, i.e., extended DPs. As such they may cliticize or be CD-ed. *La* is a spell-out of case and person or merely a spell-out of case, by further desemanticization. The clitic is an alternative realization of [Person].

2.4 The internal structure of the inflectional dative phrase

The analysis of *la*-dat suggests a parallel treatment for the dative morphology, K_{dative} , which may also be viewed as a person exponent. Just as with *la* datives, nouns which take the inflectional dative come from the lexicon endowed with an [uPerson] feature. This feature is valued KP-internally, when K_{dative} has an [*iPerson*] feature. Alternatively, if its semantic feature is bleached, K_{dative} does nothing but realize case. In such situations the KP comes out [*uPerson*, *uCase:Dative*], and clitic doubling will again be obligatory to value and spell-out Person.



At this point an important empirical remark is that outside the vP , the Inf=dat is not sensitive to AH. The distinction between [Person]-sensitive datives [=KPs], and datives which are not [Person]-sensitive [possibly DPs] is relevant for case checking. Datives which are not sensitive to [Person] are always case-licensed by prepositions, overt or null. The three overt prepositions (e.g., *grație/datorită/mulțumită* “due to”/ “thanks to”) which assign dative introduce adjuncts (adverbials of reason), as in (19):

- (19) *grație ajutorului acordat de părinți*
 thanks-to help.the.DAT granted by parents
 “thanks to the help granted by his parents”

Datives not sensitive to person are also selected by other lexical heads, e.g., adjectives like (e.g., *util* “useful”, *folositor* “useful”, *inferior* “inferior”, etc.), and they often alternate with lexical PPs, as in (20) below. Since adjectives are not case-assigners, one plausible analysis is to assume that the Inf=dat is case-licensed by a null preposition, which incorporates into the lexical head. The status of the Inf=dat phrase is again that of a PP.

- (20) *nivel de trai inferior celui din Bulgaria/ față de cel*
 level of living inferior the.DAT in Bulgaria/ with respect to that
din Bulgaria
 in Bulgaria
 “living standard inferior to that of Bulgaria”

Conclusions

Person-sensitive Inf=dat are KPs, open to CD. K_{dative} spells out case and person, or only case. Inf-dat which are not person-sensitive are ‘DPs embedded in PPs’, since they get case from prepositions.

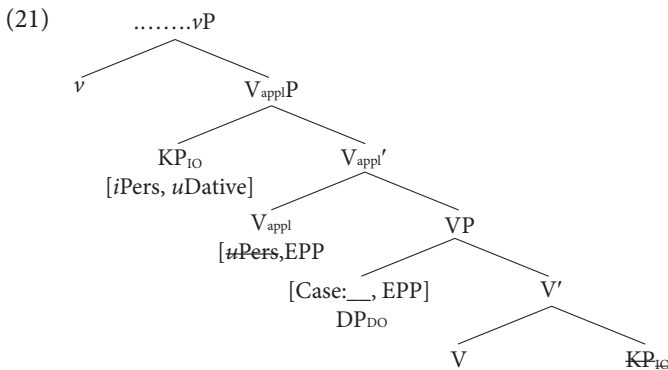
3. On the dual categorial status of datives in ditransitive constructions

3.1 Aim of the section, framework of the analysis

As far as the ν P is concerned, the categorial status of datives is less clear. Given their characteristic properties, i.e., sensitivity to [Person] and CD, we have analyzed them as extended DPs, either $KP[iPerson]$ or $KP[uPerson]$, in which case CD is obligatory.

The aim of this section is to prove that there is a class of contexts where ν P-internal datives must be PPs in order to be case-licensed. This means that ditransitives show a genuine syntactic alternation between projecting the dative as a PP or as a direct argument KP/DP . We use the same applicative derivational analysis presented in earlier work (e.g., Cornilescu et al. 2017). For reasons explained in this previous work, the basic ditransitive configuration is a Theme over Goal one; the dative is licensed by an applicative head (V_{appl}) between ν P and VP. Understandably, the features of V_{appl} are [$uPers$ ___, $uCase$:___]. Both features are **strong**. V_{appl} agrees with the dative KP , checking its case. The dative KP values V_{appl} 's [$uPerson$] feature and is finally attracted to $Spec, V_{appl}P$, since V_{appl} has strong features.

Adopting these assumptions, it is immediately apparent that the binding and scope properties of the two internal arguments do not require the projection of the dative as a PP.



Consider first sentences (1) and (2) above, where an IO asymmetrically c-commands the DO as shown by the possibility of binding a possessor inside the DO. In (1), represented in (21), the IO is undoubled, i.e., $KP [iPerson]$, in sentence (2), the IO is doubled, i.e., $KP [uPerson]$.

In derivation (21), the dative moves to a position where it is accessible to V_{appl} (a higher VP specifier, Dogget 2004). There is Agree between the dative $KP[iPerson]$

and V_{appl} . The dative checks case, values V_{appl} 's [$u\text{Person}$] feature and is attracted to $\text{Spec},V_{\text{appl}}\text{P}$. If KP is [$u\text{Person}$], the person feature remains unvalued and it will be checked in a Person Phrase at the vP periphery, as explained above.

As to binding of the IO by the DO, illustrated in (3) and (4) above, this does not require projecting the dative as a PP either. Consider example (3) where the DO binds a possessor in the IO. To obtain the desired reading nothing is required beyond configuration (21), except that in order to obtain the DO>IO reading, one must assume that at LF the IO reconstructs. A similar analysis may work when the IO is cliticized, moving even higher than $V_{\text{appl}}\text{P}$ (4). Reconstruction into the merge position remains available. Thus the basic facts confirm Pineda's (2012) claim that there is no PDC with Romance ditransitive verbs. In the next subsections we present counterexamples to this claim.

3.2 Multiple datives

A frequent case in point is that of double datives, featuring a possessive dative, necessarily cliticized, and a low Goal/Recipient dative, as in (22a):

- (22) a. *Ion și-a vândut casa unor rude/ la*
 Ion he.REFL.DAT.CL-has sold house.the some.DAT relatives at
niște rude.
 some relatives.
 "Ion sold his house to some relatives."
- b. *Ion le-a vândut casa lui unor rude/ la*
 Ion they.DAT.CL-has sold house.the his some.DAT relatives at
niște rude.
 some relatives.
 "Ion sold his house to some relatives."

In (22a), the reflexive clitic *-și* is a possessive dative clitic which binds the direct object (DO), *casa* "the house" and receives its ϕ -features by agreement with the subject. In this context, the Goal/Recipient dative cannot be CD-ed. If the Goal/Recipient is doubled the possessor must occur in its original position, as a Genitive inside the DP_{theme} , as in (22b), where there is no possessor dative clitic, while the Goal/Recipient is CD-ed. Consider the assignment of case in (22a) now. There is agreement between researchers (e.g., Pylkkänen 2002, 2008; Deal 2013) that the possessive dative construction is an applicative one, where a DP-internal possessor raises to $\text{Spec},V_{\text{appl}}\text{P}$ to check case and then cliticizes. Therefore in possessive dative sentences, the dative possessor must be licensed by V_{appl} . But if this is the case, the Goal/Recipient cannot agree with the same V_{appl} to value its case feature. The only

means of licensing the Goal is a null preposition which licenses the Goal's case and incorporates into the lexical V, as described for APs. Examples of this type, which are quite common, represent genuine Prepositional Dative Constructions.

3.3 Narrow (asymmetric) scope

Another property of low dative indirect objects (IO) in multiple dative constructions is that they take narrow scope with respect to the DO, and there tends to be rigid DO>IO order. Consider the following set of examples:

- (23) a. *Ion a arătat un tablou de Aman fiecărui oaspete.*
 Ion has shown a painting by Aman every.DAT guest
 "John showed a painting by Aman to every guest."
 b. *Ion a arătat fiecărui oaspete un tablou de Aman.*
 Ion has shown every.DAT guest a painting by Aman.
 "Ion showed every guest a painting by Aman."
- (24) a. ^{*2}*Ion și-a arătat un tablou de Aman fiecărui oaspete al său.*
 Ion he.REFL.DAT.CL-has shown a painting by Aman every.DAT guest his
 "John showed a painting of his by Aman to every guest."
 b. ^{*}*Ion și-a arătat fiecărui oaspete al său un tablou de Aman.*
 Ion REFL.DAT.CL-has shown every.DAT guest his a painting by Aman
 "John showed a painting to every guest."
 c. *Ion și-a arătat tabloul de Aman fiecărui oaspete al său.*
 John REFL-has shown painting.the by Aman every.DAT guest of his
 "John showed his painting by Aman to every guest of his."

The pair in (23) behaves as expected. In (23a), the indefinite is preferably interpreted with wide scope, given DO>IO word order, while in (23b), the existential is preferably given narrow scope, since it is overtly c-commanded by the universal quantifier.

The picture is different in (24). The DO is bound by the possessive dative clitic, suggesting a unique specific or definite referent of the DO, as shown by the fully acceptable (24c). The attempt to construct a narrow scope reading of the indefinite by changing word order (24b) leads to severe ungrammaticality. This is expected under our analysis: the low dative is a PP and has no reason to raise and no landing position either, since its scope-taking position Spec,V_{AppI}P is blocked by the possessive dative argument.

3.4 The interference with Differential Object Marking (DOM)

Romanian DOM-ed DOs are marked by the preposition *pe* ‘on’, deriving from the Old Romanian locative-directional preposition *pre*. DOM-ed DOs are sensitive to both AH and Def-H (Tigău 2011), so that personal pronouns and proper names must be DOM-ed and animate, in particular, humans may be DOM-ed. Given the data, a parallel treatment of *pe*-DO with *la*-IO as KPs is natural (see Cornilescu 2020). Nouns high in the AH come from the lexicon marked [*u*Person] and the preposition *pe* ‘on’ is a spell out of this feature, as with *la*-datives. Given that both internal arguments are marked for the same feature, locality effects are expected to occur and they do.

As seen above in (3)–(4), bare DOs unproblematically bind a possessor contained in a dative IO, whether the latter is CD-ed or not. The picture changes when the DO is DOM-ed. It is still possible for a DOM-ed DO to bind into an undoubled IO (3), but if the IO is *doubled*, the sentence is *ungrammatical* (4), unexpectedly.

- (25) DOM-ed DP_{theme} > DP_{goal}

*Comisia a repartizat pe mai mulți medici; rezidenți unor
board.the has assigned pe more many medical residents some.DAT
foști profesori de-ai lor.
former professors of theirs.*

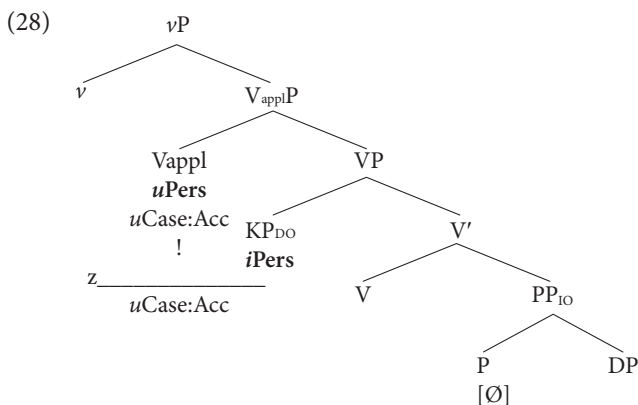
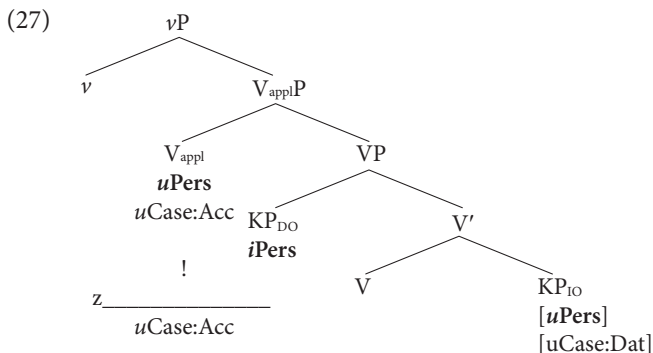
‘‘The board assigned several medical residents to some former professors of theirs.’’

- (26) *DOM-ed DP_{theme} > cl- DP_{goal}

**Comisia le-a repartizat pe mai mulți
board.the they.DAT-has assigned pe more many
medici; rezidenți unor foști profesori de-ai lor.
doctors residents some.DAT former professors of theirs*

‘‘The board assigned several medical residents to some former professors of theirs.’’

The puzzle is why there should be a difference of grammaticality between the doubled and the undoubled object. Consider the clitic case first, where the *v*P has the structure in (27). Both arguments are endowed with person and valued case features. Since the dative argument is CD-ed, its analysis as a KP is mandatory. When *V* raises to *V*_{appl} and *V*_{appl} probes its *c*-command domain, the DOM-ed object is the first that it encounters, so *V*_{appl} agrees with the ‘closer goal’ and values its own person and case features and it further attracts the KP-DO to its Specifier since its features are EPP. The IO is trapped in its merge position, and cannot check case and person anymore, so that the derivation crashes. Sentences like (26) are clear instances of intervention effects: the closer DO blocks Agree with the more remote IO.



In contrast, the derivation of (25) is quite unproblematic. Given that in this case the dative is not doubled, it will be projected as a PP (28). The null preposition assigns case and incorporates into V.

Generalizing over the two cases, the dative *must be analyzed as a PP* when case cannot be assigned by V_{appl} because the latter licenses a clitic argument. PP datives cannot be CD-ed. Furthermore they take asymmetric narrow scope, behaving like prepositional datives in the English PDC.

3.5 Non-configurational semantic effects of the Romanian dative alternation

An insightful proposal regarding the semantics of alternative syntactic constructions is Beavers & Nishida (2010), who claim that when there are alternative variants of the same thematic structure, one of the variants ‘has stronger truth conditions’, i.e., it is more informative, allowing more entailments. One parameter of strength, relevant for the (Romanian) dative alternation is the direct/oblique coding of an argument (DP > PP), combined with the presence/absence of a clitic which doubles

the DP argument. Roughly, what increases in the stronger variant is ‘the degree of affectedness’ (in the sense of Beavers 2011).

The semantic effect of the contrast between direct and prepositional coding is clearly visible in Ficiary constructions where the doubled KP dative (inflectional or *la*-dative) may alternate with a lexical prepositional construction. Consider examples like (29). Sentence (29b) has stronger truth conditions, suggesting the current existence of a dative referent, a possible direct contact of the Agent with the Beneficiary, which is thus more affected, than in the prepositional construction (29a).

- (29) a. *Ion a cumpărat jucării pentru nepoții lui.*
 John has bought toys for grandchildren his
 “Ion has bought toys for his grandchildren.”
- b. *Ion le-a cumpărat jucării nepoților lui.*
 Ion they.DAT.CL-has bought toys grandchildren.the.DAT his.
 “Ion has bought toys for his grandchildren.”
- (30) a. *Două biciclete au fost furate vecinilor.*
 two bikes have been stolen neighbours.the.DAT
 “My neighbours got two of their bikes stolen.”
- b. *Două biciclete au fost furate de la vecini.*
 two bikes have been stolen from neighbours
 “Two bikes were stolen from the neighbours/My neighbours got two of their bikes stolen.”

The pair (30a–b) also shows the relevance of KP vs PP coding. In (30b) with a PP coding of the IO, the interpretation of the PP is either Source or Maleficiary. In sentence (30a) with a dative KP IO, the interpretation is uniquely Maleficiary. In sentences (31a–b), with *la*-phrases, the clitic can make the difference between a PP Goal assigned accusative, and a dative argument, with an intended Recipient interpretation.

- (31) a. *Am trimis scrisoarea la părinți/ la poștă pentru o vecină.*
 have.I sent letter.the at parents at post-office for a neighbour.
 “I sent a letter to my parents/to the post-office for a neighbour.”
- b. *Le-am trimis scrisoarea la părinți.*
 they.DAT.CL-have.I sent letter.the at parents
 “I sent my parents the letter.”

Finally, recall that in the double dative cases discussed above, the low dative which is a PP must be interpreted with narrow scope, just as in English PDC.

4. Conclusions

Romanian possesses a genuine dative alternation, of the type recently (re)-defined by Ormazabal & Romero (2017). This is trivially true for Ficiary roles which show the alternation of a CD-ed dative phrase with lexical PP. But it is interestingly also true for Recipient/Goal datives where both the inflectional and the *la*-dative must be given a dual categorial analysis, either as person sensitive extended DPs (=KPs) or as PPs, in order to be properly licensed. The PP analysis is a last resort when V_{appl} is not accessible. Semantically the KP-dative alternant has stronger truth conditions than the PP alternant, increasing affectedness.

Acknowledgements

We are grateful to the audience of Going Romance in Bucharest, December 2017 and to the anonymous reviewers for helpful comments. We would also like to thank Adina Dragomirescu and Alexandru Nicolae for their support and editing this volume.

Funding

The research for this paper has been funded by the Alexander von Humboldt Foundation for the second author.

References

- Anagnostopoulou, Elena. 2005. Cross-linguistic and cross-categorial variation of datives. In Melita Stavrou & Arhonto Terzi (eds.), *Advances in Greek generative syntax. In honor of Dimitra Theophanopoulou-Kontou*, 61–126. Amsterdam: John Benjamins.
<https://doi.org/10.1075/la.76.05ana>
- Beavers, John. 2011. On affectedness. *Natural Language & Linguistic Theory* 29(2). 335–370.
<https://doi.org/10.1007/s11049-011-9124-6>
- Beavers, John & Chiyo Nishida. 2010. The Spanish dative alternation revisited. In Sonia Colina (ed.), *Romance linguistics 2009: Selected papers from the 39th Linguistic Symposium on Romance Languages*, 217–230. Amsterdam: John Benjamins.
<https://doi.org/10.1075/cilt.315.13bea>
- Belletti, Adriana. 2005. Extended doubling and the VP periphery. *Probus* 17(1). 1–35.
<https://doi.org/10.1515/prbs.2005.17.1.1>
- Cornilescu, Alexandra. 2020. Ditransitive constructions with DOM-ed direct objects in Romanian. In Anna Pineda & Jaume Mateu (eds.), *Dative constructions in Romance and beyond*, 117–142. Berlin: Language Science Press.

- Cornilescu, Alexandra, Anca Dinu & Alina Tigău. 2017. Romanian dative configurations: Ditransitive verbs, a tentative analysis. *Revue roumaine de linguistique* 62(2). 157–178.
- Cuervo, Cristina. 2003. Datives at large. Doctoral dissertation, MIT.
- Deal Amy Rose. 2013. Possessor raising. *Linguistic Inquiry* 44(3). 391–432.
https://doi.org/10.1162/LING_a_00133
- Diaconescu, Constanța Rodica & María Luisa Rivero. 2007. An applicative analysis of double object constructions in Romanian. *Probus* 19(2). 209–233.
<https://doi.org/10.1515/PROBUS.2007.007>
- Dogget, Teal. 2004. All things being unequal: Locality in movement. Doctoral dissertation, MIT.
- Harley, Heidi. 2002. Possession and the DOC. In Pierre Pica & Johan Rooryck (eds.), *Linguistic Variation Yearbook*, Volume 2. 31–70. Amsterdam: John Benjamins.
<https://doi.org/10.1075/livy.2.04har>
- Iorga, Ana. 2013. O tipologie a dativului românesc. Doctoral dissertation, University of Bucharest.
- Kayne, Richard. 1984. *Connectedness and binary branching*. Dordrecht: Foris.
<https://doi.org/10.1515/9783111682228>
- Longobardi, Giuseppe. 2008. Reference to the individuals, person and the variety of mapping parameters. In Henrik Høeg Müller & Alex Klinge (eds.), *Essays on nominal determination*, 189–212. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.99.11lon>
- Ormazabal, Javier & Juan Romero. 2017. Historical changes in Basque dative alternations: Evidence for a P-based (neo)derivational analysis. *Glossa* 2(1). 1–39.
<https://doi.org/10.5334/gjgl.103>
- Pineda, Anna. 2012. Double object constructions and dative/accusative alternations in Spanish and Catalan: A unified account. *Borealis* 2(1). 57–115. <https://doi.org/10.7557/1.2.1.2524>
- Pylkkänen, Liina. 2002. Introducing arguments. Doctoral dissertation, MIT.
- Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge, MA: MIT Press.
<https://doi.org/10.7551/mitpress/9780262162548.001.0001>
- Richards, Mark. 2008. Defective agree, case alternation, and the prominence of person. *Linguistische Arbeitsberichte* 86. 137–161.
- Tigău, Alina. 2011. *Syntax and interpretation of the direct object in Romance and Germanic languages with an emphasis on Romanian, German, Dutch and English*. Bucharest: Editura Universității din București.

The interpretation of null subjects in Romanian

An information-structure approach for comparative analysis

Mara Frascarelli

Università degli Studi Roma Tre

This chapter explores the ‘acceptability’ and ‘interpretation’ of referential null subjects in Romanian in different syntactic conditions and distinct clause types (matrix clauses, embedded under bridge verbs, embedded under factive verbs and adverbial clauses). Based on the results of an original online survey (carried out by almost 80 respondents), it is shown that Romanian is a consistent *pro*-drop language, in which no partial *pro*-drop properties can be found. Furthermore, the results emerged from the analysis provide significant support to the validity of an information-structure approach to the interpretation of null subjects, in which the Topic Criterion (Frascarelli 2007), the formation of Topic chains and the existence of silent Topics play a crucial role.

Keywords: null subjects, consistent/partial *pro*-drop, Topic Criterion, Topic Chain, A-Topic, G-Topic, silent Topic, locality

1. The Null Subject Parameter and the interpretation of null subjects from an IS-perspective

1.1 A short overview

In its original formulation (Perlmutter 1971), the *pro*-drop parameter aimed to capture the empirical observation that in some languages a ‘definite, referential, pronominal subject’ must be expressed in all finite clauses.¹ This observation was resumed, extended and re-elaborated in the form of the ‘Extended Projection Principle’ in Chomsky (1982) as an addendum to the former Projection Principle

1. In this work we concentrate on referential null subjects; other types such as arbitrary and expletive null subjects are left aside.

(Chomsky 1981), a basic tenet of Generative Grammar. Since then, interest in null subjects has never decreased and several works have attempted to define the formal properties that determine the setting of its different options, yielding consistent, partial, radical, semi-*pro*-drop and non-*pro*-drop languages (for background details, discussion and references, cf. Holmberg et al. 2009; Biberauer et al. 2010; Cognola & Casalicchio 2018).

Since the seminal work of Jaeggli & Safir (1989) and Rizzi (1982, 1986), the Null Subject Parameter (NSP) has been claimed to be dependent on the φ -features that are specified or encoded in the relevant licensing head (i.e., the Inflectional head node).

Assuming a novel approach, Frascarelli (2007) moves the focus of analysis from licensing to the interpretation of licensed null subjects. That is to say, assuming with Holmberg (2005) that the licensing of a null subject (hereafter, NS) depends on the presence of a D-feature in T, Frascarelli concentrates on how a licensed NS can be correctly given a referential index.²

In this respect, an information-structure (IS) strategy is proposed, according to which the interpretation of a referential NS depends on a matching relation ('Agree', in minimalist terms; cf. Chomsky 1995) between *pro* and a specific type of Topic, that is to say, the 'Aboutness-shift' Topic (A-Topic). The latter is characterized in Frascarelli & Hinterhölzl (2007) as the Topic merged in the highest Topic position in the C-domain (ShiftP; cf. (1) below), whose head is specified for the [about(ness); sh(ift)] features. This means that the A-Topic is endowed with the discourse property of proposing 'what the sentence is about' (using Reinhart's 1981 definition), also providing a 'file card' for the mental folder in which relevant information is stored.

$$(1) \quad [\text{ForceP} [\text{ShiftP}[\text{about}; \text{sh}] \text{A-Topic}_k [\text{ContrP} [\text{FocP}[\text{FamP}[\text{FinP} [\text{TP} \text{pro}_k \dots]]]]]]]]]$$

[Agree]

1.2 The Topic Criterion and the formation of Topic chains

Since every predicational (i.e., non-thetic) sentence must have a topic (cf., among others, Kuroda 1965; Lambrecht 1994; Krifka 2007) and Agree is a local relation, the matching relation illustrated in (1) must be realized for *every* occurrence of an NS.

2. As a matter of fact, φ -features only provide functional information about the person, number and gender of an NS, but they do not explain how it can be syntactically 'anchored' to the intended antecedent in a *pro*-drop language.

Nevertheless, conversation implies a ‘multi-clausal domain’, in which different propositions might ‘be about’ the same Topic. In this case, the established A-Topic is kept continuous across sentences and needs not be overtly realized in the subsequent C-domains. This means that *silent* (i.e., null) copies of an established A-Topic must be assumed.

This framework of analysis is formalized in the so-called Topic Criterion, which accounts for the IS-identification of a referential *pro*:

- (2) **Topic Criterion** (re-elaborated from Frascarelli 2007, § 5.3)
 - a. The high Topic field in the C-domain contains a position in which the [aboutness; shift] feature is encoded and matched (via Agree) by the local (3rd person) NS;
 - b. When continuous, the A-Topic can be null (i.e., silent).

Furthermore, from a conversational dynamics perspective, Bianchi & Frascarelli (henceforth, B&F 2010) have later provided evidence that the A-Topic is a *conversational move* insofar as topic selection is a ‘speech act itself’ (cf. Krifka 2001: 25). Consequently, its realization is dependent on illocutionary force and, as such, restricted to *root clauses*. This is formalized as follows:

- (3) **Interface Root Restriction (IRR; B&F 2010, § 7)**
Information Structure phenomena that affect the conversational dynamics (CG management) must occur in clauses endowed with illocutionary force that implement a conversational move.

Topic continuity thus gives rise to the creation of *Topic chains*, which must be started from root (or root-like) clauses and include overt and silent copies that provide a local antecedent to the NSs realized in the relevant chain:

- (4) **Topic Chain Condition** (Frascarelli 2018, § 9.3)
 - a. An A-Topic-chain can only be started from a root (or root-like) C-domain;
 - b. The A-Topic heading the Topic chain can be silent.

Frascarelli’s (2007) theory has been supported by comparative analyses on a number of typologically different languages and, though that proposal dealt with ‘consistent’ NS languages in its original formulation (i.e., with languages that can realize NSs in all syntactic conditions), subsequent evidence has been provided showing that the Topic Criterion (2) should be considered as a ‘macro-parameter’ and that the Topic Chain Condition (4) can account for the interpretation of NSs in ‘partial’ and ‘radical’ NS languages as well (Frascarelli 2018; Frascarelli & Jiménez-Fernández 2019; Frascarelli & Casentini 2019). In this picture, partial and radical properties rely on IS conditions, while differences depend on interface conditions (PF visibility) and independent syntactic restrictions (cf. cited works for details).

The present work contributes to this comparative analysis and evaluate the Topic Criterion in Romanian, a Romance language for which no systematic analysis has ever been dedicated to the interpretation of NSs and the creation of Topic chains. Relevant data for analysis will be drawn from the original experiment illustrated in § 2 below.

2. The experiment: Structure and methodology

2.1 Objective and working hypothesis

The present experiment basically replicates the survey originally designed and carried out in Frascarelli (2018) to compare Italian, a consistent NS language, with partial *pro*-drop in Finnish and Russian. The objective is to evaluate the acceptability and interpretation of NSs in Romanian and see whether and to what extent Romanian can be considered a consistent *pro*-drop language. For this purpose, the experiment has been translated and distributed to a large number of native informants, so as to check their judgments on a number of matrix and subordinate clauses, in which the indicative mood is always used.³ The results obtained have been thus systematically compared with the data collected for Italian, Finnish and Russian.

2.2 Structure and informants

The test was loaded on a dedicated website and distributed online (hence, participation was free and no selection was imposed). We thus collected 215 questionnaires, even though only 79 of them were completed and could be considered for analysis. Though anonymous, the questionnaire included a preliminary section asking for demographic and sociolinguistic information (age, sex, provenience, education, familiarity with linguistics), which is given in Table 1 below:

3. Hence, clauses featuring the subjunctive mood have not been tested. This decision was chiefly due to the necessity of realizing a ‘systematic comparison’ between Romanian and the languages tested in previous experiments, in which the same questionnaire has been used. Nevertheless, we reckon that this limitation does not lessen the significance of results and the validity of relevant conclusions, since the indicative is the most frequent mood, used for factual statements and positive beliefs, and it can be found in all languages. Hence, we surmise that the syntactic behaviour emerging from the present analysis can provide a reliable characterization of the *pro*-drop quality of Romanian.

Table 1. Background data on informants*

TOT	Age (av.)	Sex		Education		Field		Linguistics competence
		M	F	Univ.	Other	Human.	Other	Yes
79	32	19%	81%	93%	7%	85%	15%	74%

* As to origin, 90% of informants were resident in the Bucharest area, which can be plausibly considered as representative of the standard variety of Romanian. Of course, dialectal variation can be an interesting subject for future research (we thank an anonymous reviewer for this suggestion).

The questionnaire included 7 conditions (listed below). Each condition was tested through two target sentences and each sentence was used twice so as to analyse the interpretation of either an NS or an overt pronoun in the same clausal type, namely:

- a. in the complement of a bridge verb,
- b. in the complement of a factive verb,
- c. double embedded with an intervening 3rd person singular pronoun,
- d. having an embedded DP as intended antecedent,
- e. in a temporal adverbial clause (with and without an overt Topic in initial position),
- f. in a conditional adverbial clause,
- g. in a matrix sentence without an overt Topic.

The test thus included a total of 28 target sentences (a necessary limitation to avoid tiredness), to which 6 distractors were added, for a total of 34 tokens.⁴ Target sentences were conveniently distanced and randomized and, importantly, informants could not 'go back' in the survey (hence, they could not 'correct' or double-check their former answers). All sentences were preceded by a context introducing possible antecedents.

Acceptability was expressed on a Likert scale from 0 to 4 (with 0 being totally unacceptable, and 4 definitely acceptable).⁵ Only with a positive evaluation (i.e., either 3 or 4) would a sub-question follow, asking for the subject of the action expressed in the target sentence. Three alternatives were given and informants were asked to indicate one of them.

4. An anonymous reviewer points out that the number of fillers is rather limited. We acknowledge that fillers should be at least 50% of total sentences; nevertheless we decided to reduce their number to avoid a lengthy test, which is very likely to be left uncompleted by informants. Furthermore, we reckon that each of the seven conditions tested in the experiment could serve as a 'distractor' for the others, since informants were confronted with different constructions and phenomena and no sing-song effect could be produced.

5. We preferred to use the notion of 'acceptability' rather than 'grammaticality' since a Likert scale implies the possibility of a 'gradient' evaluation for the sentences under examination, while grammaticality requires a clear-cut judgment, which could bias the present investigation.

3. Data analysis: Results and discussion

3.1 NS in the complement of a bridge verb

The first structural context we deal with is the case of a *pro* sitting in the complement of a ‘bridge verb’. As is generally known, this is a ‘root-like’ context, insofar as it allows for the realization of a number of phenomena that are only allowed in root sentences and in a subset of root-like subordinate clauses.⁶ Importantly, one such phenomenon concerns the realization of an A-Topic in the relevant embedded C-domains (cf. B&F 2010). Let us then consider sentence (5) and the interpretations provided for NSs in Table 2 below (here and in the rest of the contribution, Romanian results are compared with the Italian, Russian and Finnish data originally presented in Frascarelli 2018):⁷

- (5) [John is telling Mary news about two common friends, Leo and his brother, since he met Leo some days before. John says:]
- Leo_k ha detto che pro ha comprato una casa* (ITA)
 - Leo a zis că pro a cumpărat o casă* (ROM)
 - Leo_k kertoi, että pro oli ostanut talon* (FIN)

6. The ‘bridge’ term traditionally refers to verbs allowing for long-distance dependencies (i.e., verbs of saying and opinion, like *say* and *think*, whose complement clauses do not form a ‘barrier’ for movement). Hence, the most used syntactic diagnostic to evaluate verbs as belonging to this group is checking whether they can be subject to long movement operations, like in (i):

- (i) [*What_k do you think* [_{tk} *that Mary said* [_{tk} *that Ann gave John* [_{tk} *for his birthday party*]]]]?

However, providing a comprehensive list of bridge verbs is not easy. In fact, a very debated topic in the literature is exactly whether a verb should join the bridge class or not for its semantic properties. For example, it seems to be generally agreed that verbs with a more factive interpretation (Kiparsky & Kiparsky 1974) permit extraction more readily, so that predicates of stating and thinking are usually better than verbs of asking. Furthermore, frequency also plays a role: if a verb is sufficiently frequent it becomes more transparent to movement.

Given the above, the following list can be provided for Romanian, since these verbs are generally considered bridge verbs ‘cross-linguistically’: *a spune* (“say”), *a afirma* (“assert”), *a explica* (“explain”), *a crede* (“believe”), *a gândi* (“think”), *a nega* (“deny”), *a admite* (“admit”), *a confirma* (“confirm”), *a conveni* (“agree”).

In studies dedicated to root phenomena, complements of bridge verbs play a major role since they have a ‘quasi-root’ character, that is, they allow for the realization of root operations, while this is not possible in complements of factive or volitional verbs (cf., among others, B&F 2010; Gärtner 2001; Haegeman 2002; Emonds 2004; Meinunger 2004; Heycock 2006).

7. Note that the ‘*pro*’ notation was not used in the survey, to avoid confusions with the informants (possibly not familiar with linguistic terms). However, it is used here to illustrate the relevant examples.

- d. *Lev_k skazal čto pro kupil dom* (RUS)
 “Leo_k said that he_k/his brother_z bought a house.”⁸

Table 2. NS embedded under a matrix bridge verb

	ACCEPT	Leo		His brother		Both are possible	
		Nr	%	Nr	%	Nr	%
ITA	128/128	31	24%	40	31%	57	45%
ROM	79/79	26	33%	26	33%	27	33%
FIN	273/273	188	69%	38	14%	47	17%
RUS	53/53	26	50%	7	12%	20	38%

As shown, this clausal type allows for embedded NSs in all the languages under examination, with interpretive differences, though. In particular, while in partial NS languages (Finnish and Russian) the matrix Topic-‘subject’ is more frequently selected for the antecedent role, no preference can be detected in Romanian (like in Italian), in which answers are equally distributed between the three options, showing that, in the absence of the prosodic cues that characterize an A-Topic, ambiguity dominates (consistent with Frascarelli’s 2018 conclusions).

These results also show that an NS is *not necessarily subject-oriented* in consistent *pro-drop* languages (differently from what is claimed in Filiaci et al. 2014) and that the interpretation of an NS does not depend on overt ‘syntactic control’.⁹ In this respect, statistical analysis shows that Romanian and Italian are equivalent, since the difference between the informants who selected *Leo* and those who selected *his brother* is not significant according to the Fisher Exact Test ($p = 0.5837$), as well as the difference between the informants who judged both options as possible and those who selected either *Leo* or *his brother* ($p = 0.1144$ and $p = 0.3924$, respectively).

Let us now consider the case in which an overt pronoun is realized in the embedded clause, to see whether any difference can emerge with respect to NS interpretation. Consider (6) below and the data in Table 3 (the context is the same as in (5) above):

- (6) a. *Leo_k ha detto che lui ha comprato una casa*
 b. *Leo a zis că el a cumpărat o casă*
 c. *Leo_k kertoi, että hän oli ostanut talon*
 d. *Lev_k skazal čto on kupil dom*
 “Leo_k said that he_k/his brother_z bought a house.”

8. Note that the four languages under examination will be always illustrated in the order given in (5).

9. Also note that *his brother* is a possible antecedent in Finnish and in Russian as well. Hence, partial *pro-drop* languages also show some form of ‘gradience’ in this respect (on this notion, see the *Conclusions*).

Table 3. Overt pronoun embedded under a matrix bridge verb

	ACCEPT	Leo		His brother		Both are possible	
		Nr	%	Nr	%	Nr	%
ITA	128/128	35	27%	36	28%	57	45%
ROM	79/79	32	40%	32	40%	15	20%
FIN	273/273	60	22%	22	8%	191	70%
RUS	53/53	20	38%	3	6%	30	56%

Table 3 shows that in Romanian, like Italian, overt pronouns are *not discourse-context oriented* and that, as emerged for NSs, no preference is attested for either exclusive referential interpretation. As a matter of fact, the difference between the values recorded for *Leo* and those recorded for *his brother* is not significant ($p = 1.0000$), in both languages.

Nevertheless, in this case Romanian shows a significant difference with respect to Italian concerning the selection of the ‘both’ option: while this answer is dominant in Italian (and the other languages examined), a specific referential option is preferred in Romanian (whereas any option is equally fine with an NS, cf. Table 2 above). Indeed, the difference between the values scored for the ‘both’ option in Italian and in Romanian with respect to either *Leo* and *his brother* are statistically significant (i.e., $p = 0.0011$ and $p = 0.0013$, respectively).¹⁰

This difference seems to show that in Romanian overt pronouns trigger a ‘focus’ interpretation, possibly connected with a ‘corrective effect’, which makes the ‘both’ option (almost) immaterial. However, since no context is provided in the experiment, this interpretation only relies on the discourse role assigned to the matrix subject (*Leo*) by informants: if it is considered an A-Topic, coreference with a focused pronoun is excluded (since A-Topics can only form Topic chains), if it is seen as a background element, coreference with *lui* can be triggered.¹¹

Based on the data just discussed, we can conclude that, in the absence of interface (prosodic) cues, Romanian (like Italian) equally admits as a possible antecedent for a *pro*(noun) either an overt A-Topic (i.e., the Topic-‘subject’ in the matrix

10. On the other hand, it should be noted that, when an exclusive referential reading is at stake, partial *pro*-drop languages decidedly select an ‘overt’ Topic as an external antecedent (*Leo*, in this case), a preference that is statistically clearly significant in Finnish $p < 0.0001$ and very significant in Russian ($p = 0.0010$). The interface requirement for an overt antecedent seems to be a specific condition of partial NS languages (Frascarelli & Jiménez-Fernández 2019). The fact that an overt antecedent is not preferred over a silent one in Romanian provides additional evidence that this language has no partial *pro*-drop qualities.

11. Though interesting, a thorough treatment of this case is beyond present purposes and left open for future research on the interaction between antecedent selection of overt pronouns and (embedded) correction. We thank an anonymous reviewer for this suggestion.

clause,¹² cf. (7a) below) or a silent A-Topic, referring to an entity that is part of the context but is not overtly realized in the relevant matrix C-domain (as in (7b)). According to the framework assumed, Romanian thus qualifies as a consistent *pro*-drop language.

- (7) a. [_{ShiftP} *Leo*_k [_{TP} *pro*_k *a zis* [_{ForceP} [_{Force} *că* [_{ShiftP} <*Leo*>_k [_{TP} *pro*_k / *el*_k *a cumpărat o casă*]]]]]]]
 b. [_{ShiftP} <*his brother*>_k [_{TP} *Leo a zis* [_{ForceP} [_{Force} *că* [_{ShiftP} <*his brother*>_k [_{TP} *pro*_k / *el*_k *a cumpărat o casă*]]]]]]]

3.2 NS in the complement of a factive verb

Let us now examine the interpretation of an NS that is embedded under a factive verb. This condition is particularly interesting since the complement of a factive verb is presupposed information (cf. Meinunger 2004) and, as such, it does not have root-like properties.¹³ This means that the C-domain of a clause embedded under a factive verb cannot host an A-Topic (contrary to the complements of bridge verbs, cf. (7a–b) above), but only G-Topics,¹⁴ which can be linked to a silent A-Topic in the matrix C-domain.

12. In this respect, it is important to remember that referential subjects in NS languages are assumed to be sitting in an A'-position (cf. Frascarelli 2007); this means that a preverbal DP like *Leo* in (5)–(6) is in fact a Topic (while an NS is merged in the canonical subject position). Nevertheless, since the '*pro*' notation was not used in the survey (cf. fn. 7), matrix Topics appear as preverbal 'subjects' in the target sentences.

13. As in the case of bridge verbs, a commonly agreed definition and an exhaustive list of factive verbs are hard to provide (for important discussion, cf. Meinunger 2004). Indeed, semantic properties play a central role also in the definition of this verb class, which generally includes emotive and 'truly' factive verbs (contrary to bridge verbs, no assertion is made when true factives are used; cf. Kiparsky & Kiparsky 1974). What is expressed is a fact, an emotion or a state of mind (i.e., something that cannot be negated). Commonly assessed factive verbs are a *displăcea* ("resent"), a *regreta* ("regret"), a *ii părea rău* ("be sorry"), a *fi surprins* ("be surprised"), a *deranja* ("bother"), a *deplânge* ("deplore"), a *fi jignit* ("take offence").

Syntactic heuristic for this verb type is the block imposed on long movement operations and root phenomena, like VP preposing, as shown respectively in (i)–(ii) below:

- (i) * [*What*_k *do you regret* [*t_k that Mary said* [*t_k that you might want* *t_k for your birthday party*]]]?

(ii) * *I resent that never in my life did I do something like that.*

14. The G-Topic can be considered as a D-linked constituent, either in a 'strong' (Heim 1982) or in a 'weak'/'familiar' sense (Roberts 2003). According to Frascarelli & Hinterhölzl's (2007) typology, (at least) two types of G-Topics should be distinguished: (i) 'Aboutness' G-Topics, which are part of a Topic Chain and, as such, they serve a 'continuity' function as 'low copies' of an established A-Topic; (ii) 'Background' G-Topics, which are not linked to the current A-Topic and serve to retrieve given/ background information.

The data concerning one level of embedding will not be discussed since they basically show the same results illustrated for bridge verbs in § 3.1 above, and we will immediately consider the interpretation of an NS in complex sentences with two controlling antecedents (condition (c) in § 2.2):

- (8) [Jari is going with Leo to the race]
- Jari_ksi dispiace che Leo_z pensa che **pro** perderà la gara*
 - Lui Jari îi pare rău că Leo crede că **pro** va pierde concursul.*
 - Jaria_k harmittaa, että Leo_z ajattelee, että **pro** häviää kilpailun.*
 - Iari_k¹⁵žal čto Lev_z думаet čto **pro** proigaet corebnovanje.*
“Jari_k is sorry that Leo_z thinks that he_k/he_z will lose the race.”

Table 4. NS double embedded under a matrix factive verb

	ACCEPT	Jari		Leo		Both are possible	
		Nr	%	Nr	%	Nr	%
ITA	128/128	20	16%	45	35%	63	49%
ROM	79/79	21	27%	45	57%	13	16%
FIN	273/273	32	12%	167	61%	74	27%
RUS	53/53	5	10%	28	53%	20	37%

As we can see, the embedded antecedent (*Leo*) is selected by the majority of Romanian informants, with a difference that is statistically significant with respect to either the matrix subject (*Jari*; $p = 0,0180$) or the ‘both’ option ($p = 0.0003$), as in partial NS languages. Hence, in this respect, Romanian is crucially different from Italian.

Since the Topic-‘subject’ embedded under a factive verb cannot be an A-Topic (but only a G-Topic), this result leads to the conclusion that the relevant chain is headed by a *silent A-Topic*, which is embedded under the bridge verb, as shown in (9a) below:

- (9) a. [_{TP} *lui Jari îi pare rău* [_{ForcpePcă} [_{FamP} *Leo_k* [_{backgr}]_{TP} *t_kcrede* [_{ForcpePcă} [_{ShiftP} <Leo_k> [_{TP} *pro_kva pierde concursul*]]]]]]]]

In other words, the DP *Leo*, which is background information in the first embedded clause, is (silently) proposed for a topic shift after the bridge verb.

The preference for this interpretation seems to show that Romanian is sensitive to a ‘locality’ requirement, so that the *closest* link is more frequently selected for

15. The ‘intended’ subject of the factive predicate ‘be sorry’ in Russian is marked with a DAT Case.

antecedence in a Topic chain (like in partial *pro*-drop languages). On the other hand, this requirement does not seem to be operative in a language like Italian, in which no significant difference is attested between the equivalent of sentence (9a) and a Topic chain that is headed by a silent A-Topic in the matrix C-domain (as in (9b) below):

- b. [ShiftP <Jari_z> [TP *lui Iari_z îi pare rău* [Forccep*că* [FamP *Leo_k* [backgr] [TP *t_k crede* [Forccep*că* [FamP <Jari_z> [about] [TP *pro_z va pierde concursul*]]]]]]]]]¹⁶

We will return to and discuss in more detail this locality requirement and its consequences in § 4.

Finally, when an overt pronoun is present in the relevant target sentence, like in (10) below, co-reference with the matrix Topic-‘subject’ is preferred (47%) with respect to the closest DP *Leo* (32%).¹⁷ Even though this difference is not statistically significant ($p = 0.1833$), we can take it as the sign of a slight obviation tendency for overt pronouns in Romanian, so that the furthest referent is preferred.

- (10) *Lui Iari îi pare rău că Leo crede că el va pierde concursul.*

3.3 NS in adverbial clauses

It is generally agreed that adverbial clauses are not endowed with illocutionary force; hence, they cannot host A-Topics according to the IRR (3). Furthermore, Haegeman (2004, 2012) argues for an important distinction between central and peripheral adverbial clauses according to which, while the left periphery of central adverbials totally lacks the functional projections encoding speaker-related functions (speech time, epistemic modality, illocutionary Force) and are within the scope of operators, peripheral adverbial clauses seem to admit the realization of (some) root phenomena.

Since we were interested in examining the interpretation of NSs in non-root adverbial clauses (i.e., in structural contexts that do not allow for an A-Topic in the embedded C-domain), we limited our survey to central adverbial clauses, considering in particular temporal and conditional clauses (cf. conditions (e) and (f) in

16. Note that Aboutness G-Topics are merged in Spec,FamP since they are always part of a Topic chain. On the other hand, Background G-Topics are merged in Spec,TP and can be either moved to the C-domain or realized in situ, in cross-linguistic variation (cf. Jiménez-Fernández & Miyagawa 2014).

17. See Frascarelli (2018) for the relevant table with comparative details.

§ 2.2. above).¹⁸ Specifically, central clauses were tested in two structural conditions: with the adverbial clause in a post-matrix position (which is considered as ‘basic’ in Haegeman 2012) and with the adverbial clause in a pre-matrix position (which is usually defined as ‘fronted’), so as to check the interpretive effect of a supposed movement operation.

3.3.1 Conditional clauses

When the conditional clause is located in post-matrix position, the NS was accepted by 100% of informants in all the languages examined (including partial *pro*-drop ones). Consider the target sentence in (11) and the interpretive data shown in Table 5 below:

- (11) [Pedro’s friends meet for a beer. They know that Leo is still at work with his boss. They hope he can join them later. One of them says:]
- Leo può venire se pro finisce il lavoro*
 - Leo poate să vină dacă pro termină treaba*
 - Leo voi tulla jos pro saa työn tehtyä*
 - Lev možeť přijti, esli pro zakačivaet raboty*
“Leo_k can come if (he_k/his boss) finishes the work.”

Table 5. NS embedded in a post-matrix conditional clause

	ACCEPT	Leo		His boss		Both are possible	
		Nr	%	Nr	%	Nr	%
ITA	128/128	84	66%	16	12%	28	22%
ROM	79/79	48	61%	19	24%	12	15%
FIN	273/273	210	77%	17	6%	46	17%
RUS	53/53	30	56%	13	25%	10	19%

18. An anonymous reviewer points out the importance of showing, through relevant diagnostics, that temporal and conditional clauses in Romanian behave like in English, in order to validate their use in the present experiment. We thank the reviewer for this suggestion but, for space reasons, we will limit this comparison to a single phenomenon, namely ‘argument fronting’. As shown below, this phenomenon is blocked in Romanian as in English, proving that conditional adverbial clauses resist root phenomena, as is expected for central adverbial clauses (relevant examples have been provided and double-checked by native informants):

- *Am crezut că în timp ce eu acest ziar citeam tu merseseși la bancă.*
 - *I thought that while this paper I was reading you had gone to the bank.*
- *I-am spus lui John că dacă aceste examene nu voi trece nu putem merge în vacanță vara viitoare.*
 - *I told John that if these exams I won’t pass we cannot go on vacation next summer.*

For further discussion on the semantic, discourse and interface characterization of central and peripheral clauses, see Frascarelli (2019).

As shown, the values scored by the matrix preverbal DP (*Leo*) in adverbial clauses are dominant, with a difference with respect to those attested with bridge verbs (cf. Table 2) that is extremely significant in Italian and Finnish ($p < 0.0001$) and significant in Romanian ($p = 0.0472$) and Russian ($p = 0.0304$). This means that the possibility of establishing *his boss* as a silent head for the Topic chain is very low in all the languages examined, irrespective of their *pro*-drop quality. This result strongly supports the non-availability of A-Topics in central adverbial clauses and the validity of the present IS-approach to the interpretation of NSs.

Let us now consider the values attested for the pre-matrix position in Romanian, compared to consistent Italian and partial NS languages (the context being the same as in (11) above):

- (12) a. *Se pro finisce il lavoro, Leo_k può venire.*
 b. *Dacă pro termină treaba Leo poate să vină.*
 c. *Jos pro saa työn tehtyä, Leo_k voi tulla.*
 d. *Eslī pro zakačivaet raboty, Lev_k možet prijti.*
 “If (*he_k/his boss_z*) finishes the work, *Leo_k* can come.”

Table 6. NS embedded in a pre-matrix conditional clause

	ACCEPT	Leo		His boss		Both are possible	
		Nr	%	Nr	%	Nr	%
ITA	128/128	84	66%	16	12%	28	22%
ROM	79/79	30	38%	35	44%	14	18%
FIN	131/273	34	26%	63	48%	34	26%
RUS	25/53	7	28%	14	56%	4	16%

As we can see, while in partial *pro*-drop languages an NS located in a pre-matrix adverbial clause is only accepted by (circa) 50% of informants, this syntactic condition is totally fine for Romanian and Italian informants.

As for interpretation, no significant difference can be attested in Romanian between an external and a clause-internal antecedent. This result is in line with the analysis provided for the interpretation of NSs embedded under bridge verbs and, as such, fully expected: in both cases a silent A-Topic is the *closest link* for the NS. This means that the embedded DP *Leo* can be either interpreted as an Aboutness G-Topic (13a), which maintains the chain started by the silent A-Topic, or as a Background G-Topic (13b), without significant preferences:¹⁹

19. In Italian, on the other hand, the DP *Leo* in the matrix clause is mostly interpreted as an Aboutness G-Topic (interpretation (13a)), in line with the results attested for the post-matrix position (cf. Table 5 above). However, no structural or interface reasons can be found for such a tendency, which should be therefore ascribed to a general preference for a ‘continuing function of G-Topics in Italian’ (a topic for future research, possibly based on corpus data).

- (13) a. [_{ShiftP} <Leo_i> [_{CP} *se pro_i finisce il lavoro*] [_{FamP} *Leo_k* [_{IP} *pro_k può venire*]]
↑
[about; sh] [about] G-Topic
- b. [_{ShiftP} <*his boss*> [_{CP} *dacă pro_i termină treaba*] [_{FamP} *Leo_k* [_{IP} *t_k poate să vină*]]
↑
[about; sh] [backgr] G-Topic

Note further that the interpretive difference attested for partial *pro*-drop languages (preferring an external referential reading, as in (13b)) shows that in Romanian and Italian the preverbal position of conditional adverbial clauses is not derived, hence no trace is left within the matrix clause and no intervenience effect (cf. Haegeman 2012) can be determined by the movement of the DP *Leo*. This supports the hypothesis that pre-matrix adverbial clauses in Romance languages can be considered as ‘Frames’, merged in a dedicated position in the C-domain (cf. Frascarelli 2017). This is not apparently the case in languages like Finnish and Russian, in which the pre-matrix position is also characterized by low acceptance values (cf. Frascarelli 2018 for discussion).

Since the results obtained with conditional clauses are very consistent across the two adverbial types, in the following section we will discuss temporal clauses for different conditions, in order to provide additional data and arguments for the present proposal.

3.3.2 Temporal clauses

As indicated in condition (e) in § 2.2, we wanted to check whether and to what extent the presence of an overt A-Topic, which can only be interpreted as the subject of the adverbial clause, can be accepted as a chain head. A sample sentence for this condition is given in (14):²⁰

- (14) a. **Anna, mentre pro va a scuola, Maria mangia una mela.*
 b. **Ana, în timp ce pro se duce la școală, Maria mănâncă un măr.*
 c. **Anna, kun pro menee kouluun, Marja syö omenan.*
 “Anna_k, when (she_k/she_z) is going to school, Mary_z eats an apple”

Interestingly, this sentence was utterly rejected by informants in the three languages examined and, specifically, by 91% of Italian, 87% of Romanian and 96% of Finnish speakers.

In the present approach, this result can find an immediate explanation, since the initial DP *Anna* cannot be an antecedent for the relevant NS. More precisely, (i) if located in the adverbial C-domain, the relevant DP cannot be an A-Topic (cf.

20. This condition was not tested in the Russian experiment.

IRR (3)), (ii) if located in the matrix C-domain, it could not be associated with any syntactic role in the matrix sentence and, finally, (iii) if considered a Hanging Topic, it cannot serve as an antecedent for *pro* (as is argued in Frascarelli 2007). In conclusion, this result and relevant considerations strongly support the Topic Criterion (2) and the Topic Chain condition (4), with particular reference to (a) the dedicated IS-function of A-Topics as antecedents for *pro* and (b) their exhaustive location in C-domains that are endowed with illocutionary force.

Additional support to the present argumentation is given by the fact that a sentence like (15) is fully accepted (100%) by both Italian and Romanian speakers, while the corresponding Finnish example (15c) only scored 43%:

- (15) a. *Maria, mentre Anna va a scuola, pro mangia una mela.*
 b. *Maria, în timp ce Ana se duce la școală, pro mănâncă un măr.*
 c. *Maria, kun Ana menee kouluun, pro syö omenan.*
 “Mary_k, when Anna_z is going to school, (she_k/she_z) eats an apple”

This result is predicted by the present approach because the DP *Maria* can be interpreted as an A-Topic connected with the matrix NS. On the other hand, this condition obtains marginal effects in a partial *pro*-drop language like Finnish, due to the pre-matrix position of the adverbial clause, which creates intervention effects (see Frascarelli 2018 on this point).

3.4 Embedded DP as an intended antecedent

Let us now consider the judgments provided for sentences designed to check the necessity (or, at least, the relevance) of syntactic control for the licensing and interpretation of an NS in Romanian (that is to say, condition (d) in § 2.2).

Consider the following sentence, in which a DP modifier (i.e., a DP embedded in a complement PP) is proposed as a non c-commanding antecedent:

- (16) a. *Il discorso di Leo ha chiarito che pro non è colpevole.*
 b. *Discursul lui Leo a clarificat că pro nu este vinovat.*
 c. *Leo-n puhe tek-i selvä-ksi, ett-ei pro ole syyllinen.*
 “Leo_k’s talk made it clear that (he_k/sb. else) was not guilty.”

This type of sentence was accepted by Romanian and Italian informants (while it is very marginal in Finnish), and the interpretation provided bluntly excludes the necessity of overt syntactic control. This provides additional support to the Topic Criterion (2) and, in particular, to the necessity of assuming silent A-Topics. Consider Table 7 below:

Table 7. NS with a non-commanding covert antecedent

	ACCEPT	Leo		Sb else		Both are possible	
		Nr	%	Nr	%	Nr	%
ITA	128/128	26	20%	24	19%	78	61%
ROM	79/79	24	30%	33	42%	22	28%
FIN	131/273	47	36%	14	11%	70	53%

The values attested for the non *c*-commanding DP *Leo* in Romanian are roughly the same as those obtained for the *c*-commanding preverbal subject DP in bridge constructions (cf. Table 2 above). This provides substantial evidence that the licensing and interpretation of an NS in *pro*-drop languages does not depend on syntactic control.

3.5 NS embedded in clauses without an overt Topic in the matrix C-domain

The data examined in the previous sections consistently supported the existence of silent A-Topics, in line with the Topic Criterion (2). As a consequence, when an NS is realized in a matrix clause, a silent A-Topic must be assumed in the relevant C-domain. This is illustrated in (17) for Italian and represented in (17'):

- (17) *Vorrei presentarti Leo_k. È il mio*
 want.COND.1SG introduce.INF-2SG.IOCL Leo be.3SG the my
miglior amico.
 best friend
 ‘I’d like to introduce Leo_k to you. (he_k) is my best friend.’

- (17') *Vorrei presentarti Leo.*
 [_{ShiftP} <Leo_k > [_{TP} *pro_k è il mio miglior amico*]]
 [-----]

So far, Romanian has shown the properties of a consistent *pro*-drop language, with a (significant) structural preference for the closest antecedent (due to a locality requirement; cf. § 3.2). Given the above, the realization of an NS in a matrix clause that does not contain an overt A-Topic in its-C-domain is expected to be fully accepted in Romanian.

Let us now consider the results obtained with sentences in which a matrix NS (with no Topic) is preceded by a sentence including two plausible overt antecedents, with different syntactic functions. The first case is one in which the choice is between two argument DPs:

- (18) a. *Jari ha parlato a Leo ieri.*
Ora pro ha capito cosa è successo.
 b. *Jari a vorbit cu Leo ieri.*
Acum pro a înțeles ce s-a întâmplat.
 c. *Jari puhui Leolle eilen.*
Nyt pro ymmärtää, mi-tä tapahtu-i.
 d. *Jari pogovoril co Lvom bčera.*
Teper' pro ponjal, čto proizošlo
 “**Jari_k** talked to **Leo_z** yesterday. Now (**he_k**/**he_z**) understood what happened.”

Table 8. NS in a matrix clause hosting no overt Topic

	ACCEPT	Jari		(to) Leo		Both are possible	
		Nr	%	Nr	%	Nr	%
ITA	128/128	67	53%	19	15%	42	32%
ROM	79/79	31	39%	26	33%	22	28%
FIN	87/273	23	26%	12	14%	52	60%
RUS	33/53	13	39%	4	13%	16	48%

The results obtained show that silent A-Topics are fully accepted as antecedents for a matrix *pro* in Romanian, as expected. Furthermore, when it comes to interpretation, Romanian informants do not show any significant preference for the syntactic function of the overt DPs in the preceding clause, showing that syntactic functions are not relevant for the interpretation of an NS, like in Italian.

Let us turn now to a case in which the selection is between an argument and a non-argument DP, while the NS is embedded under a matrix clause whose C-domain contains no overt Topic, as in (19) below:

- (19) a. *Jari è andato al cinema con Leo.*
So che pro era molto contento.
 b. *Jari a fost la cinematograful cu Leo.*
Știu că pro era foarte bucuros.
 c. *Jari meni elokuviiin Leon kanssa*
Tiedän, että pro oli oikein iloinen.
 d. *Jari pošol b kino co Lvom.*
Ja znaju, čto pro byl očen' rad.
 “**Jari_k** went to the cinema with **Leo_z**. I know that (**he_k**/**he_z**) was very happy.”

Table 9. NS in an embedded clause hosting no overt Topic

	ACCEPT	Jari		(with) Leo		Both are possible	
		Nr	%	Nr	%	Nr	%
ITA	128/128	61	48%	7	6%	60	46%
ROM	79/79	31	39%	26	33%	22	28%
FIN	73/273	6	9%	14	18%	53	73%
RUS	21/53	5	26%	2	9%	14	65%

As we can see, while this sentence type determines a decrease of acceptance in partial NS languages, due to the syntactic embedding of the NS, Romanian informants totally accept this structural condition (as is the case of Italian speakers), providing further evidence for its consistent *pro*-drop quality. As for interpretation, Romanian shows *exactly* the same percentage values attested in the previous case (cf. Table 8 above), confirming the non-relevance of syntactic functions for antecedence selection.

To conclude this analysis, let us briefly focus on consistent *pro*-drop languages and check whether a ‘by-subj’ can be selected as an antecedent in Romanian. This possibility has been attested for Italian (cf. Frascarelli 2018), even though it was argued to be totally excluded in previous work (cf. Samek-Lodovici 1996). Consider sentence (20) below:

- (20) a. *Il convegno è stato presentato da Marco.*
Poi pro è andato a fare lezione.
 b. *Seminarul a fost prezentat de către Marco.*
După aceea pro s-a dus să țină cursul.
 “The conference was presented by Marco_k. Then (he_k/sb. else) went to hold his class.”

Table 10. by-SUBJ as an antecedent for NS

	ACCEPT	(by) Marco		Sb. else		Both are possible	
		Nr	%	Nr	%	Nr	%
ITA	93/128 (73%)	93	100%	0	0%	0	0%
ROM	61/79 (78%)	57	94%	4	6%	0	0%

As we can see, this is the only case in which acceptance is not full in Romanian (like in Italian), though definitely above marginality. This shows that a syntactic ‘block’ cannot be assumed (pace Samek-Lodovici 1996) but, rather, an interpretive restriction can be proposed, which is operative (for some speakers) at the interface between discourse and syntax.

Let us finally see whether an overt Topic, which is not connected with the subject position, can interfere and/or be preferred as an antecedent for an NS. A sample sentence is given in (21):

- (21) a. *A Leo non ha ancora parlato Marco:*
pro è sempre così occupato!
 b. *Cu Leo, Marco încă nu a vorbit:*
pro este mereu atât de ocupat!
 “Marco_k has not talked yet to Leo_z: (he_k/he_z) is always so busy!”

Table 11. NS in a matrix clause hosting a non-subject Topic

	ACCEPT	Marco		(a) Leo (Topic)		both are possible	
		Nr	%	Nr	%	Nr	%
ITA	128/128	60	47%	35	27%	33	26%
ROM	79/79	25	32%	34	43%	20	25%

As shown, Italian and Romanian show a different behaviour in this case. Specifically, Italian speakers prefer to link the NS to a silent A-Topic that is coreferent with the (postverbal) subject in the previous sentence (cf. (22a)) and this preference is significant with respect to the selection of either the Topic PP ($p = 0.0305$) or the ‘both’ answer ($p = 0.0208$). On the other hand, Romanian informants take the DP embedded in the overt Topic PP in the previous clause (*cu Leo*) as the most feasible discourse connection (as in (22b)), even though no significant differences can be attested with the other two options:

- (22) a. *A Leo non ha ancora parlato Marco.*
 [ShiftP <Marco>_k [TP *pro_k è sempre così occupato*]].
 b. *Cu Leo, Marco încă nu a vorbit.*
 [ShiftP <Leo>_k [TP *pro_k este mereu atât de ocupat*]].

This difference might be connected to a preference for discourse-driven continuity in Romanian, as well as to word order and its connection to IS, a working hypothesis to be explored in future research.

To conclude, the data examined clearly attest Romanian as a consistent *pro*-drop language, in which the selection of the Topic heading the chain is not driven by syntactic functions and is only dependent on IS conditions.

The preference for the closest A-Topic can be attributed to an independent structural condition, whose cross-linguistic validity has been attested for a number of phenomena, that is to say, the Minimal Link Condition (cf. Rizzi 1990; Manzini 1992 and subsequent works). Specifically, we propose that this condition can be operative in *pro*-drop languages as a strategy to reduce ambiguity when more than one preverbal DP qualifies as a plausible A-Topic from a syntax-discourse perspective and no prosodic cues are available (as in written texts), in cross-linguistic variation.²¹ On the other hand, since intonation can identify the different Topic types, this strategy is expected to be immaterial in spoken data – a prediction to be checked in future research.

To conclude, the present experiment has provided evidence that *degrees of partiality*²² arise depending on two specific variables, namely (i) overt vs silent A-Topics and (ii) local vs non-local links. This result is put forth as a proposal to be used in future research, since these variables can be taken as diagnostics to test (degrees of) partiality in NS languages, supporting Chomsky's (1995) idea that syntax feeds interpretation.

Furthermore, since the syntactic conditions used in this experiment have been selected to check the crucial role of A-Topics as antecedents, we can also conclude that the Topic Criterion (2) and the creation of Topic Chains are cross-linguistic requirements for the interpretation of *pro* in NS languages, thus supporting the theory that the acceptability and interpretation of NSs essentially depend on an IS strategy.

In this respect, Romanian shows the properties of a consistent *pro*-drop language in all respects insofar as, like Italian and different from partial *pro*-drop languages, it allows for silent A-Topics in all the syntactic conditions examined (cf. §§ 3.1, 3.2, 3.3, 3.4) and no connection has been attested between A-Topics and syntactic functions (cf. § 3.5).

Moreover, in both Italian and Romanian antecedent selection is neither subject-driven nor dependent on syntactic control (cf. § 3.2): either the overt matrix A-Topic-‘subject’ or a silent A-Topic qualify as possible antecedents. Finally, like in Italian, NSs are fully accepted in adverbial clauses, independently of their position (§ 3.3.1).

However, a significant difference with respect to Italian emerged in antecedent selection, insofar as Romanian speakers tend to link NSs to the *closest plausible link* in ambiguous contexts (cf. §§ 3.1, 3.2), a preference that we have interpreted in the

21. As a matter of fact, this strategy is not adopted by Italian speakers, who are not apparently ‘disturbed’ by ambiguous readings and select the ‘both’ answer more often than Romanian speakers.

22. On ‘gradience’ in grammar, see Fanselow et al. (2006); for a formal approach to gradient judgments, see Villata et al. (2016).

light of a Locality requirement, even though it can be also accredited to the lack of prosodic cues in the present experiment. Further research is needed to clarify this issue.

Finally, the results obtained provide substantial support for the theoretical framework adopted, since it has been shown that (a) Topic chains are invariably headed by preverbal DPs located in root or quasi-root contexts; (b) Background G-Topics do not interfere in Topic chains; (c) silent Topics are necessary to account for the interpretation of NSs in different syntactic conditions; (d) antecedent selection does not depend on syntactic control.

Acknowledgements

I wish to thank Atessa Theodora Duman for her invaluable precious contribution as a native speaker to the Romanian reduplication of this experiment and her collaboration in the collection of data and elaboration of relevant tables. I am also grateful to the members of the 31st Going Romance Conference (Bucharest, December 7th–9th 2017) for their feedback and suggestions. General disclaimers apply.

References

- Bianchi, Valentina & Mara Frascarelli. 2010. Is topic a root phenomenon? *Iberia* 2. 43–88.
- Biberauer, Theresa, Anders Holmberg, Ian Roberts & Michelle Sheehan. eds. 2010. *Parametric variation: Null subjects in minimalist theory*. Cambridge: Cambridge University Press.
- Chomsky, Noam. 1981. *Lectures on government and binding*. Dordrecht: Foris.
- Chomsky, Noam. 1982. *Some concepts and consequences of the theory of government and binding*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Cognola, Federica & Jan Casalicchio. eds. 2018. *Null subjects in generative grammar: A synchronic and diachronic perspective*. Oxford: Oxford University Press.
<https://doi.org/10.1093/oso/97801988815853.001.0001>
- Emonds, Joseph. 2004. Unspecified categories as the key to root constructions. In David Adger, Cécile de Cat & Georges Tsoulas (eds.), *Peripheries: Syntactic edges and their effects*, 75–120. Dordrecht: Kluwer. https://doi.org/10.1007/1-4020-1910-6_4
- Fanselow, Gisbert, Caroline Fery, Matthias Schlesewsky & Ralf Vogel (eds.). 2006. *Gradience in grammar: Generative perspectives*. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199274796.001.0001>
- Filiaci, Francesca, Antonella Sorace & Manuel Carreiras. 2014. Anaphoric biases and null overt subjects in Italian and Spanish: A cross-linguistic comparison. *Language and Cognitive Processes* 29(7). 825–843.
- Frascarelli, Mara. 2007. Subjects, topics and the interpretation of referential *pro*. An interface approach to the linking of (null) pronouns. *Natural Language & Linguistic Theory* 25. 691–734.
<https://doi.org/10.1007/s11049-007-9025-x>

- Frascarelli, Mara. 2017. Dislocations and Framings. In Elisabeth Stark & Andreas Dufter (eds.), *Manual of Romance Morphosyntax and Syntax*, 472–501. Berlin: Mouton.
<https://doi.org/10.1515/9783110377088-013>
- Frascarelli, Mara. 2018. The interpretation of *pro* in consistent and partial ns languages: A comparative interface analysis. In Federica Cognola & Jan Casalicchio (eds.), *Null subjects in generative grammar: A synchronic and diachronic perspective*, 211–239. Oxford: Oxford University Press.
- Frascarelli, Mara. 2019. Topics, conversational dynamics and the root/non-root distinction: Adverbial clauses at the discourse-syntax interface. In Valéria Molnár, Verner Egerland & Susanne Winkler (eds.), *Architecture of topics*, 139–170. Berlin: Mouton De Gruyter.
<https://doi.org/10.1515/9781501504488-005>
- Frascarelli, Mara & Marco Casentini. 2019. The interpretation of null subjects in a radical *pro*-drop language: Topic chains and discourse semantic requirements in Chinese. *Studies in Chinese Linguistics* 40. 1. Online. <https://doi.org/10.2478/scl-2019-0001>
- Frascarelli, Mara & Roland Hinterhölzl. 2007. Types of topics in German and Italian. In Kerstin Schwabe & Susanne Winkler (eds.), *On information structure, meaning and form*, 87–116. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.100.07fra>
- Frascarelli, Mara & Angel Jiménez-Fernández. 2019. Understanding partiality in *pro*-drop languages: An information-structure approach. *Syntax* 22. 162–198.
<https://doi.org/10.1111/synt.12184>
- Gärtner, Hans-Martin. 2001. Are there V2 relative clauses in German?. *Journal of Comparative Germanic Linguistics* 3. 97–141. <https://doi.org/10.1023/A:1011432819119>
- Haegeman, Liliane. 2002. Anchoring to speaker, adverbial clauses and the structure of CP. *Georgetown University Working Papers in Theoretical Linguistics* 2. 117–180.
- Haegeman, Liliane. 2004. The syntax of adverbial clauses and its consequences for topicalization. *Antwerp Papers in Linguistics* 107. 61–90.
- Haegeman, Liliane. 2012. *Adverbial clauses, main clause phenomena, and the composition of the left periphery*. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199858774.001.0001>
- Heim, Irene. 1982. The semantics of definite and indefinite noun phrases. Doctoral dissertation, University of Massachusetts Amherst.
- Heycock, Caroline. 2006. Embedded root phenomena. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, Volume 2, 174–209. Malden, MA: Blackwell.
<https://doi.org/10.1002/9780470996591.ch23>
- Holmberg, Anders. 2005. Is there a little *pro*? Evidence from Finnish. *Linguistic Inquiry* 36. 535–564. <https://doi.org/10.1162/002438905774464322>
- Holmberg, Anders, Aarti Nayudu & Michelle Sheehan. 2009. Three partial null-subject languages: A comparison of Brazilian Portuguese, Finnish and Marathi. *Studia Linguistica* 63(1). 59–97. <https://doi.org/10.1111/j.1467-9582.2008.01154.x>
- Jaeggli, Osvaldo & Ken Safir (eds.). 1989. *The null subject parameter*. Dordrecht: Kluwer.
<https://doi.org/10.1007/978-94-009-2540-3>
- Jiménez-Fernández, Ángel & Shigeru Miyagawa. 2014. A feature-inheritance approach to root phenomena and parametric variation. *Lingua* 145. 275–302.
<https://doi.org/10.1016/j.lingua.2014.04.008>
- Kiparsky, Paul & Carol Kiparsky. 1974. Fact. In Danny D. Steinberg & Leon A. Jakobovits (eds.), *Semantics: An interdisciplinary reader in Philosophy, Linguistics and Psychology*, 345–370. Cambridge: Cambridge University Press.

- Krifka, Manfred. 2001. Quantifying into question acts. *Natural Language Semantics* 9. 1–40.
<https://doi.org/10.1023/A:1017903702063>
- Krifka, Manfred. 2007. Basic notions of information structure. In Caroline Féry & Manfred Krifka (eds.), *Interdisciplinary studies on information structure*, 13–55. Potsdam: Universitätsverlag Potsdam.
- Kuroda, Sige-Yuki. 1965. Generative grammatical studies in the Japanese language. Doctoral dissertation, MIT. [Reprinted, New York: Garland Press. 1979].
- Lambrech, Knud. 1994. *Information structure and sentence form: Topic, focus and the mental representation of discourse referents*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511620607>
- Manzini, Maria Rita. 1992. *Locality: A theory and some of its empirical consequences*. Cambridge, MA: MIT Press.
- Meinunger, André. 2004. Verb position, verbal mood and the anchoring (potential) of sentences. In Horst Lohnstein & Susanne Trissler (eds.), *The syntax and semantics of the left periphery*, 313–341. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110912111.313>
- Perlmutter, David. 1971. *Deep and surface structure constraints in syntax*. New York: Holt, Rinehart and Wilson.
- Reinhart, Tanya. 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27. 53–94.
- Rizzi, Luigi. 1982. *Issues in Italian syntax*. Dordrecht: Foris. <https://doi.org/10.1515/9783110883718>
- Rizzi, Luigi. 1986. Null objects in Italian and the theory of *pro*. *Linguistic Inquiry* 17. 501–558.
- Rizzi, Luigi. 1990. *Relativized minimality*. Cambridge, MA: MIT Press.
- Roberts, Craig. 2003. Uniqueness in definite noun phrases. *Linguistics and Philosophy* 26. 287–350.
<https://doi.org/10.1023/A:1024157132393>
- Samek-Lodovici, Vieri. 1996. Constraints on subjects: An optimality theoretic analysis. Doctoral dissertation, Rutgers University.
- Villata, Sandra, Luigi Rizzi & Julie Franck. 2016. Intervention effects and relativized minimality: New experimental evidence from graded judgments. *Lingua* 179. 76–96.
<https://doi.org/10.1016/j.lingua.2016.03.004>

Verum focus and Romanian polar questions

Ion Giurgea

'Torgu Jordan – Alexandru Rosetti' Institute of Linguistics

Most previous research on the intonation of Romanian polar questions has claimed that the neutral pattern has the main prominence (nuclear accent) on the finite verb, and a final contour characteristic of questions with early focus. I argue that this pattern is not neutral, at least in information-seeking questions, but indicates verum focus; the neutral pattern has the nuclear accent on the last prosodic word, as expected given the general prosodic properties of Romanian. The data presented will also shed some light on the conditions of use of verum focus in questions.

Keywords: verum focus, polar questions, Romanian

1. Previous research on the intonation of polar questions in Romanian

Previous studies on the intonation of polar questions in Romanian (Avram 1973; Dascălu 1979a; b; Dascălu-Jinga 2001; Grice et al. 2000; Ladd 2008; Jitcă et al. 2015) agree on the fact that the nuclear accent is low, representable as L* in the ToBI notation, and the boundary tone depends on whether the nuclear accent is the last lexical accent in the sentence, or not.

- i. If the nuclear accent is not the last lexical accent – a pattern which I call the 'early-focus pattern' (EF) – the low tone attained on the nuclear accent is followed by a low plateau until the last lexical accent. On this syllable, the tone rises, yielding an accent which can be described as L+H*; then, if the last accented syllable is also the last syllable in the sentence, nothing more happens, the boundary tone being thus H% (see Example 1 and Figure 1); if it is followed by one or more unstressed syllables, the tone falls again, the boundary tone being L% (see Example 2 and Figure 2). This pattern has been analyzed by Grice et al. (2000) and Ladd (2008) as reflecting an underlying high-low boundary

tone (HL%) realized by ‘secondary association’ of the H-part with the last accented syllable (which yields the (L+)H* L% contour found in ex. 2), followed by truncation of the L part, in case the last accented syllable is sentence-final (yielding the (L+)H* H% pattern illustrated in ex. 1).

- ii. If the nuclear accent coincides with the last lexical accent – a pattern which I will call the ‘final focus-pattern’ (FF) – the boundary tone which follows the L* nuclear accent is always high (H%). Therefore, if the last accented syllable is sentence-final, the final rising takes place during this syllable (see Example 3 and Figure 3). Otherwise, the rising starts after the accented syllable, see ex. 4 and Figure 4.

- (1) [MA'RIA]_{Foc} a ve'nit?
 Maria has come
 “Is it Maria who has come?”

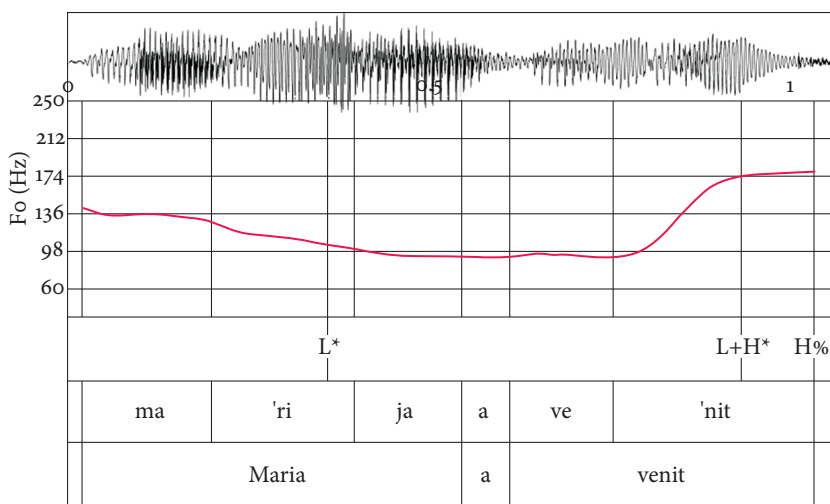


Figure 1.

- (2) [MA'RIA]_{Foc} 'vine?
 Maria comes
 “Is it Maria who’s coming?”

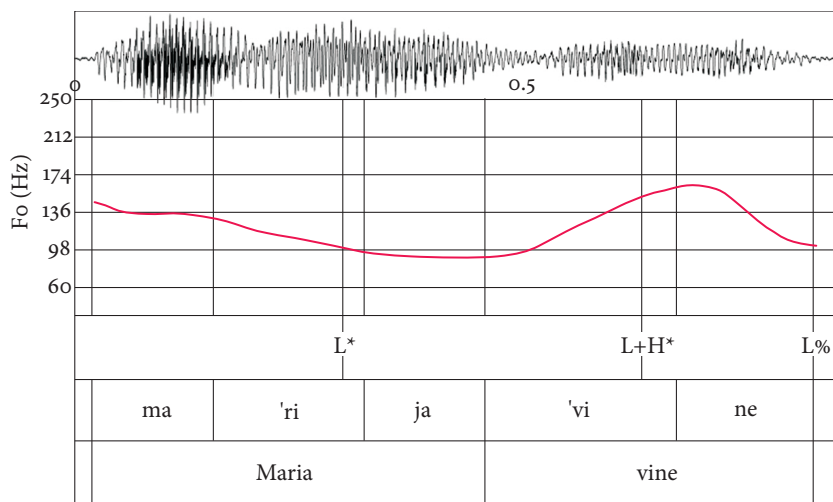


Figure 2.

- (3) Ma'ria [a VE'NIT]_{Foc}?
 Maria has come
 “Maria, did she come?” / “Did Maria come?”

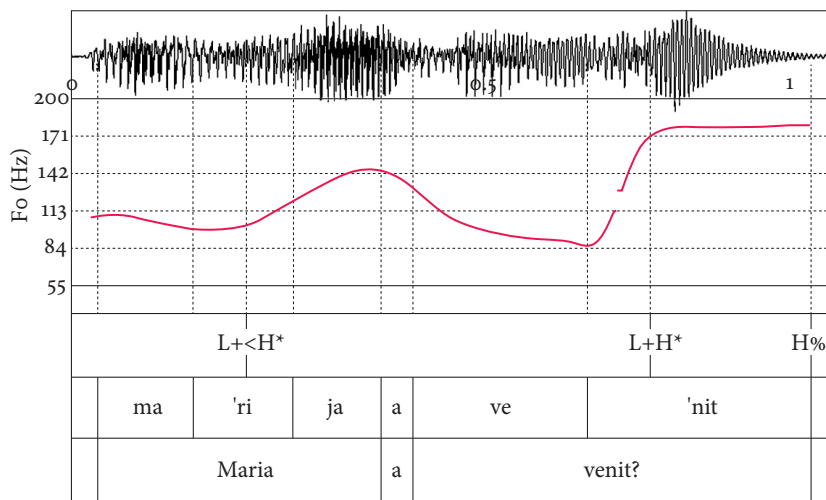


Figure 3.

(4) *Ma'ria* ['VINE]_{Foc?}

Maria comes

“Maria, is she coming?” / “Is Maria coming?”

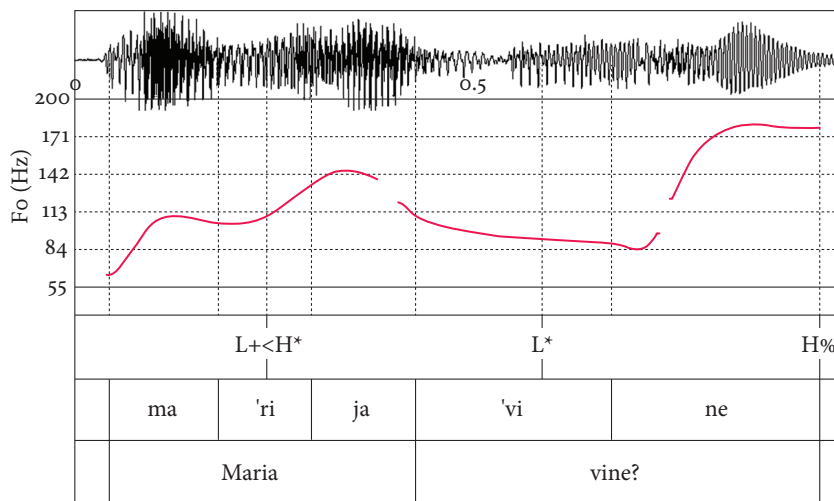


Figure 4.

The first study which described these patterns, Avram (1973), did not characterize any of these patterns as unmarked.¹ Further studies (Dascălu 1979a; b; Dascălu-Jinga 2001: 33) claimed that polar questions always contain an ‘emphasized element’, and this element is, by default, the verb, a description which is also adopted by the latest academic grammar of Romanian (see GALR II: 33, 957–958). As Romanian is not verb-final, this means that the neutral contour would be an EF contour with the nuclear accent on V. This characterization was formalized in the autosegmental-metrical model by Grice et al. (2000), who claim that the same EF pattern with main stress on the verb represents the neutral pattern in Hungarian and Greek. This claim found its way into the 2008 edition of Ladd’s classic handbook of intonational phonology.

Jitcă et al. (2015), the chapter devoted to Romanian in Frota & Prieto’s handbook of intonation in Romance, find both the EF pattern with focus on the verb and the FF pattern as neutral, based on the example reproduced in (5), where the two patterns in (5a) and (5b) were produced by two different speakers from Iași (a third speaker, from Transylvania, produced a pattern which is different from those described here; in this study, I confine myself to the intonation of the southern

1. The main objective of that study was to compare the intonations of the standard language with those of the Muscel variety, where the patterns are totally different – the main prominence is high rather than low and the final contour is always rising.

variety of standard Romanian, as spoken in Bucharest, which is similar, with respect to questions, to the Moldavian variety; the fact that Transylvania uses different patterns in questions is well-known, see, e.g., Dascălu-Jinga 2001).

- (5) [Context: somebody enters a shop]

A.'veʃi mar.me.'la.dă?

have.2PL marmalade

“Do you have marmalade?”

a. $L^* L+;H^* L\%$ (EF pattern)

b. $L^* L^* H\%$ (FF pattern)

2. The neutral pattern of information-seeking polar questions

The claim that the neutral pattern involves nuclear stress on the verb followed by deaccentuation of all the following prosodic words (i.e., the EF pattern with focus on the verb) is rather surprising, suggesting that polar questions in Romanian cannot be focally neutral (or ‘all-focus’), but must involve some narrow focus, which by default falls on the verb. One might think that polar questions must involve at least focus on polarity, phonologically realized as focus on the finite verb – Holmberg (2013) indeed proposed that polar question formation involves movement of a polarity operator to the left-peripheral Spec,FocP. However, Bianchi & Cruschina (2016) convincingly argued that, at least in Italian, polar question formation is independent of focalization, because focus fronting is possible in polar questions, and the same holds for Romanian (see Giurgea 2016 for focus fronting in polar questions in Romanian).

I argue that in fact the neutral pattern in polar questions in Romanian is identical to the FF pattern (as in (5b)), as expected given the general prosodic properties of Romanian, where default prominence is on the right (the general pattern of well-described European languages, such as Western Germanic and the rest of Romance). What looks like an EF pattern in (5a) may be due to a special intonational pattern used in polite requests with an interrogative form (see § 5 below). Here I concentrate on the information-seeking reading of polar questions, which qualifies, of course, as pragmatically neutral, reflecting the literal meaning.

Example (6) illustrates an out-of-the-blue information-seeking polar question. As illustrated in Figure 5, this sentence has the FF pattern (the L^* nuclear accent falls on the last lexical accent, on the syllable *ta-* from *tata*, and the boundary tone is H%).

- (6) [*Ce s-a întâmplat,*] *a ve.'nit 'ta.ta?*

what REFL-has happened has come father-the

“What happened? Has dad come?”

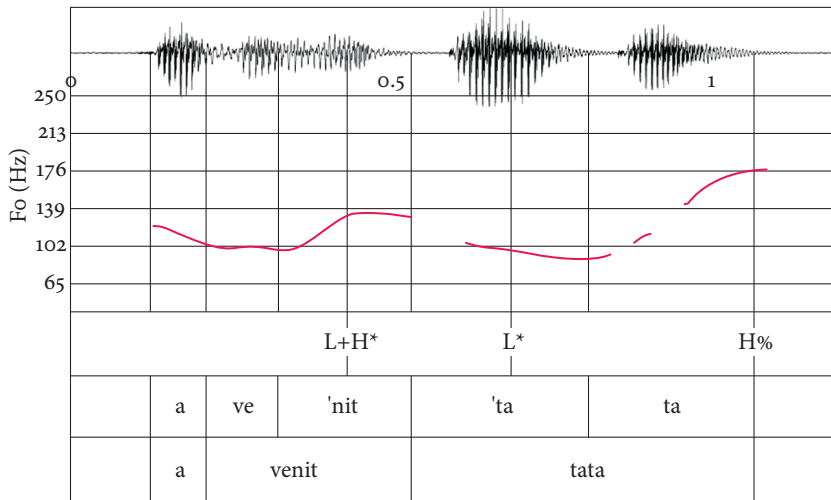


Figure 5.

The EF pattern is possible in the beginning of a conversation, but requires that the issue *whether p* has previously occurred between the conversation participants, in previous conversations. Thus, the sentence in (6) can have the EF pattern with nuclear stress on the verb in the context indicated in (7) – see Figure 6.

- (7) [Context: their father's arrival has occurred as an issue in the communicative exchanges between the conversation participants at some point before, e.g., three hours before somebody said that we have to wait for dad to get home in order to watch a DVD]

A *ve.'nit 'ta.ta?*

has come father-the

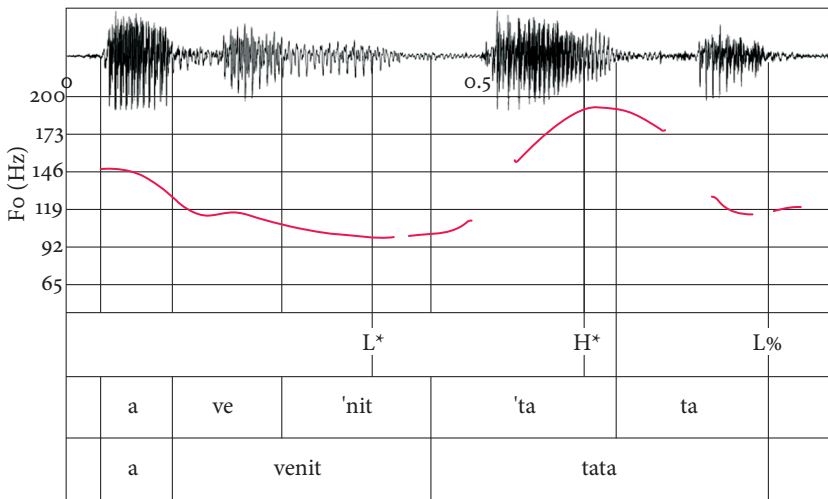


Figure 6.

More generally, the pattern with focal stress on the verb is possible if the issue *whether p* has already occurred in the linguistic exchanges between the conversation participants. This recalls the characterization of ‘verum’ in polar questions given by Gutzmann & Castroviejo Miró (2011):

- (8) ‘verum’ is only used if ?*p* is present among the questions under discussion (QUD)

I conclude that, in polar questions that are genuine requests for information, the early focus pattern with nuclear stress on the verb indicates verum focus.

3. On ‘verum focus’

‘Verum focus’, a term proposed by Höhle (1988, 1992), refers to a type of focalization manifested by sentence stress on an element that fills a (dedicated) clausal functional head position, C or Infl (which extends to the Spec of a null C, for German),² which, for declaratives, can be roughly characterized as emphasizing the assertion – see (9), where I use an idiomatic expression (*to take care* and its Romanian equivalent *a avea grijă*, [lit. ‘to have care’]) in order to distinguish, in English and Romanian, verum focus from predicate focus:

- (9) a. *But I DID take care.*
 b. *Dar am AVUT grijă.* (Ro.)
 but have.1SG had care
 c. *Ich HABE doch aufgepasst.* (Ge.)
 I have.1SG but taken-care

In English and German, the distinction between verum focus and focus on the predicate is also visible in sentences with auxiliaries (see (9c) for German), but in Romanian auxiliaries are clitics, which arguably form a complex head with the verb (see Dobrovie-Sorin 1994; Barbu 1999; Dobrovie-Sorin & Galves 2000, among others). Therefore, it is the lexical V which bears the stress in comparable situations, see (9b). To illustrate the contrast between verum focus and predicate focus we can use either idioms, as in (9b), or light verbs such as the copula, see (10):

- (10) [Context: people are wondering whether the competition has been announced.]
ESTE anunțat concursul.
 is announced competition.the
 ‘The competition IS announced.’

2. In German embedded interrogatives, verum focus can be realized by stress on the *wh*-pronoun that occupies Spec,CP.

As we have seen in (8) for polar questions, *verum focus* on a sentence expressing the proposition *p* is typically used if the issue *whether p* is present in the context (cf. also Büring 2006) – see the examples below, which illustrate *verum focus* in all major types of sentences (declarative, interrogative and directive):

- (11) [Context: people are debating on whether somebody's actions were legal or not]
 AVEA *voie*
 had.3SG permission
 "He WAS allowed to do that."
- (12) [A: *Ne așteptam toți să fie supărată* "We all expected her to be upset."
 B: *Și, ERA supărată?*
 and was.3SG upset
 "So, WAS she upset?"]
- (13) [Context: X said he was unsure whether he should allow his daughter to play outside; addressed to X:]
 DĂ-i *voie!*
 give.IMPV-3SG.DAT permission
 "DO allow her!"

In declaratives, the focus can be treated as focus on assertion (Höhle 1992) or on polarity (cf. the notion of 'polarity focus', used by Halliday 1967; Dik et al. 1981; Gussenhoven 1984, for English, and Watters 1979, for the Bantu language Aghem), but this analysis cannot extend to interrogatives (see (12)) or directives (see (13)). Therefore, Lohnstein (2012, 2016) proposed that *verum focus* is focus on what he calls 'sentence mood', which is what makes a sentence assertive, interrogative or imperative – he does not refer to illocutionary force because *verum focus* can also be found in embedded clauses.

Other studies (Romero & Han 2004; Romero 2005; Gutzmann & Castroviejo Miró 2011; Repp 2013; Gutzmann et al. manuscript) argued that 'verum' prosody signals the presence of a special operator 'verum', rather than marking focus on an independently existing element. This led Gutzmann & Castroviejo Miró (2011) to divorce 'verum' from focus completely. (14) shows their definition of 'verum':

- (14) 'Verum' is a conversational operator (with 'use-conditional meaning') which takes as input a proposition *p* and conveys that *?p* should be downgraded from the Question Under Discussion. (Gutzmann & Castroviejo Miró 2011: 144)

An evaluation of all these analyses is beyond the scope of this chapter. I would simply suggest that the 'verum focus' prosody may cover more than a single phenomenon. For instance, in some cases, such as the examples of *verum focus* in embedded contexts, we may indeed be dealing with focus on polarity – see Höhle's example in (15), and a similar Romanian example in (16):

- (15) *Da stehen die Leute die du NICHT getroffen hast. Aber dort stehen*
 there stand the people that you not met have now there stand
die Leute, DIE du getroffen hast.
 the people that you met have (Höhle 1992: 134)
- (16) *Mi-ai spus doar ce n-am voie. Acum aş*
 me-have.2SG told only what not-have.1SG permission now would.1SG
vrea să ştiu ce AM voie să fac.
 want SBJV know.1SG what have.1SG permission SBJV do.1SG
 “You only told me what I’m not allowed to. Now I’d like to know what I AM
 allowed to do.”

But there is evidence that a focus on polarity analysis cannot cover all the cases. In particular, it cannot cover verum focus in questions. Let us consider Example (12). If verum focus were focus on polarity, we should be able to build a negative version of (12), with the same meaning associated with focus on negation. But we cannot do this: the focus must occur *above* negation, as shown in the English version in (17a). In Romanian, if we put the focus on negation, as in (17b), we do not obtain a neutral sentence like (12), but a biased question, with an incredulity meaning, which indicates that something more is at play than mere presence of an antecedent for *p*, ruling out verum focus (see (12b)). Stress placement on the verb as in (17c) is completely out, showing that verum focus must scope above negation.³ But as negation obligatorily precedes the finite verb in Romanian, an order corresponding to the English (17a) is impossible. Therefore verum focus cannot be realized on the finite verb at all in this context. A possible paraphrase of (17a) is given in (17d), using an epistemic adverb under focus:

- (17) [A: “We all expected her not to be upset.” – Ro.: *Ne aşteptam toţi să nu fie supărată*]
 a. B: “So, WAS she not upset?”
 b. **Şi, NU era supărată?* (Ro.)
 and not was.3SG upset.FSG
 “# So, wasn’t she upset?”
 c. **Şi, nu ERA supărată?*
 and not was.3SG upset.FSG
 d. *Şi, aşa era? Într-adevăr nu era supărată?*
 and so was really not was upset.FSG

Previous studies on verum focus in Romanian polar questions (Giurgea & Remberger 2012, 2014) have identified two types: a ‘neutral’ or ‘informational’ verum focus type,

3. Cf. Alboiu et al. (2015), who argue for verum focus being placed in Spec,FocP, a projection higher than NegP.

used in contexts where the issue *whether p* has already been raised and left unsolved (see (18)) and a ‘mirative’ type, associated with unexpectedness of the proposition *p* with respect to its focal alternative $\neg p$, which, combined with the question meaning, leads to an incredulity meaning (a negatively biased question), see (19). The authors use verum focus in order to account for VS orders in polar questions found with types of predications where the V-initial order cannot be neutral (see (20)).

- (18) *N-am înțeles până la urmă: ARE fratele tău mașină?*
not-have.1SG understood until at end has brother.the your car
“In the end, I didn’t understand. Does you brother own a car, (or doesn’t he)?”
- (19) *E nu zău. ARE fratele tău mașină?*
INTERJ not really has brother.the your car
“No kidding! Your brother really owns a car?”
- (20) *(Fratele tău) are (#fratele tău) mașină (#fratele tău)*
brother.the your has brother.the your car brother.the your
“Your brother owns a car.”

Now, looking at prosody, even though (19) might be argued to have an emphatic stress on the V, it does not look like what we have described in § 1 as an EFF pattern, because V is not followed by a plateau until the final accent (the phrase *fratele tău* has H* on its last accent), and the final H is also less marked than in genuine EFF patterns – see Figure 7:

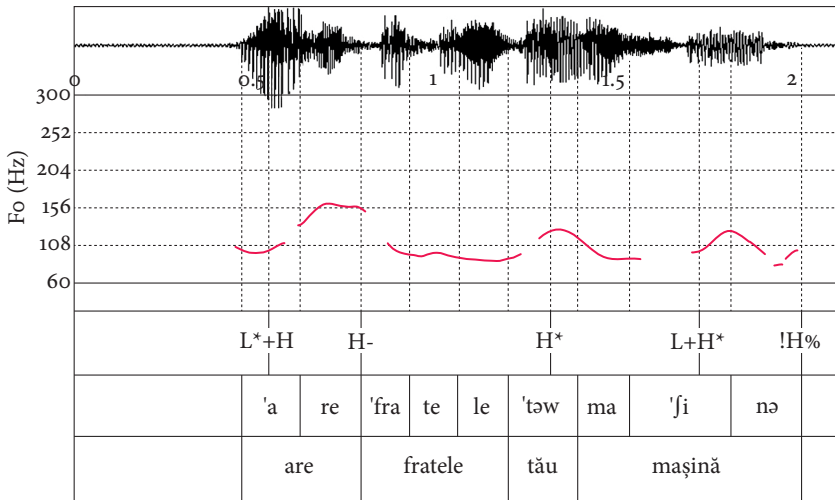


Figure 7.

A genuine EFF pattern is found in Example (18), see Figure 8:

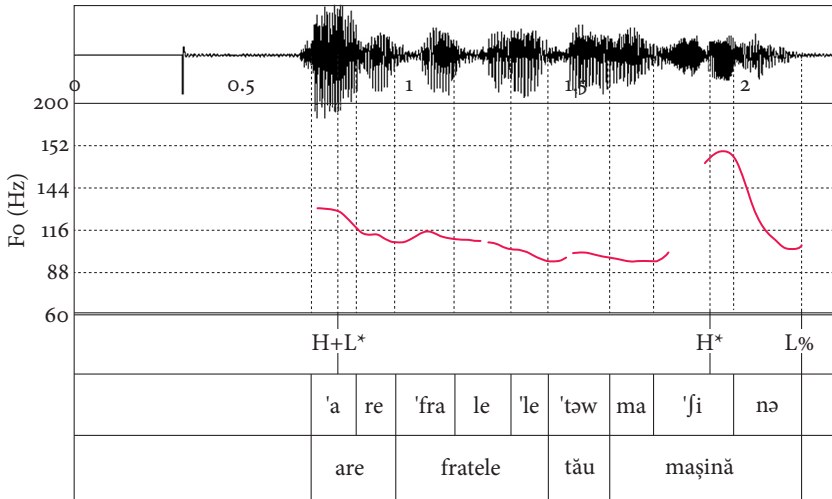


Figure 8.

Jitcă et al. (2015) also noticed that questions expressing surprise have a special pattern, with repeated L^*+H accents, as in (21):⁴

- (21) *'Spui că 'Mario candi'dează ca pri'mar?*
 say.2SG that Mario runs-as-a-candidate as mayor
 $L+<H^*$ L^*+H $H-$ L^*+H $H-$ L^* $H\%$
 “Are you saying that Mario is running for mayor?”

(Jitcă et al. 2015: 304 figure 8.18)

Given these differences, although I do not exclude that VS in (19) relies on verum focus, I will confine the discussion to non-biased examples of the type in (18), which clearly rely on verum focus.

Summing up, I have claimed that unbiased information-seeking polar questions have the nuclear stress at the end in out-of-the-blue context. Conversely, I have shown that the nuclear stress on V, which had been interpreted as reflecting the neutral pattern in the previous literature, is actually used in contexts where the issue *whether p* is contextually given, reflecting verum focus. Furthermore, I have argued that in order for this issue to be contextually given, mentioning in the immediately previous discourse is not necessary: all that is required is that the issue had occurred at some point in the recent conversational exchanges of the speaker and addressee. Since these claims are new, I have tested them experimentally. The next section discusses the results.

4. A distinct intonation for mirative focus in general has also been reported for Italian (cf. Bianchi et al. 2016).

4. Experimental evidence for the claims about the neutral intonation of Romanian polar questions

For the experimental tests, I used two pairs of examples which contain the same polar question in different (imagined) contexts, one that allows *verum focus*, due to the presence of the issue *whether p* in previous conversations (Examples (23) and (25)), and one that does not, being clearly out-of-the-blue (Examples (22) and (24)). Since for the final contour of polar questions in Romanian it is crucial whether the last accented syllable is sentence-final or not (see § 1), I tested one pair of examples where this syllable (boldfaced) is sentence-final (see (22)–(23)) and one where this syllable is not sentence-final (see (24)–(25)). In the examples below, I have boldfaced the syllable which is expected to bear the nuclear stress according to the claims made in this contribution:

- (22) [Context: you meet some acquaintances on the street by chance]
Ce faceți? Ați fost la cum.pă.ră.'turi?
 what do.2PL have.2PL been at shopping.PL
 “What’s up? Have you been shopping?”
- (23) [Context: some people were supposed to do the shopping; you check whether they fulfilled their commitment]
Gata? Ați fost la cum.pă.ră.'turi?
 ready have.2PL been at shopping.PL
 “Ready? Have you been shopping?”
- (24) [Context: you enter a room where somebody has just finished a phone conversation]
Ce faci? Ai vor.'bit cu co.'pi.ii?
 what do.2SG have.2SG talked with children.the
 “What’s up? Have you spoken with the children?”
- (25) [Context: somebody told you a few hours ago that she would try to phone her children, who are abroad]
E, ai reușit? Ai vor.'bit cu co.'pi.ii?
 hey have.2SG succeeded have.2SG talked with children.the
 “Well now, have you succeeded? Have you spoken with your children?”

I worked with 13 consultants, all living in Bucharest, aged between 29 and 69. They were asked to read the sentence aloud, as they would pronounce it in the indicated context. The experiment contained six other sentences, as fillers. The recordings have been analyzed with Praat (Boersma & Weenink 2012).

The results can be grouped into the following types (where (26) summarizes the results for the pair in (22)–(23) and (27) – those for the pair in (24)–(25)):

(26) *Ați 'fost la cum.pă.ră.'turi?*

L+<H* L* LH% Pattern A

L+H* L* LH% Pattern A'

L* L+H* H% Pattern B

(27) *Ai vor.'bit cu co.'pi.ii?*

L+<H* L* LH% Pattern C

L+<H* H+L* LH% Pattern C'

L* H* L% Pattern D

L+<H* L+H* L% Pattern E

The difference between A and A' concern the realization of the prenuclear accent, which often has a delayed maximum, being analyzed as L+<H* in Jitcă et al. (2015), a view which I have adopted here. Pattern C differs from C' by the presence of a peak before the L* on the nucleus. I consider these differences irrelevant from a phonological point of view. In keeping with the characterizations provided in § 1, patterns A–A' and C–C' reflect the FF pattern (with nuclear L* on the last accented word and final rising), whereas B and D reflect the EF pattern with L* nuclear stress on the verb and final H* L% in case of non-final last accent (see D) and L+H* H% in case of final lexical accent (see C). Note that in (26), since the final syllable is in both cases rising, the most conspicuous difference concerns the first accent, which is prenuclear in A–A' (the FF pattern) and nuclear in B (the EF pattern). As for E, it does not fit into either of the patterns. It only occurred once, and seems to be accidental, reflecting a slip from the FF pattern (see the rising, prenuclear accent on the verb) to the EF pattern (see the final configuration). Examples of the four main patterns are given in Figures 9–12.

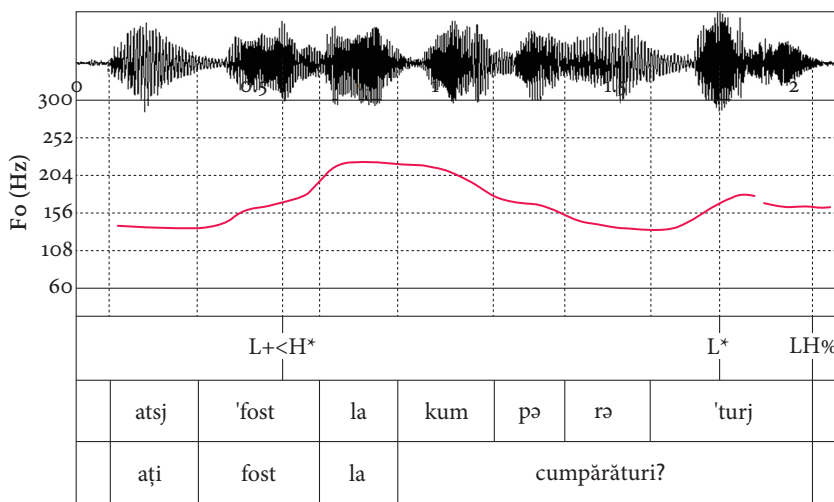


Figure 9. The A-pattern (FF, final syllable accented)

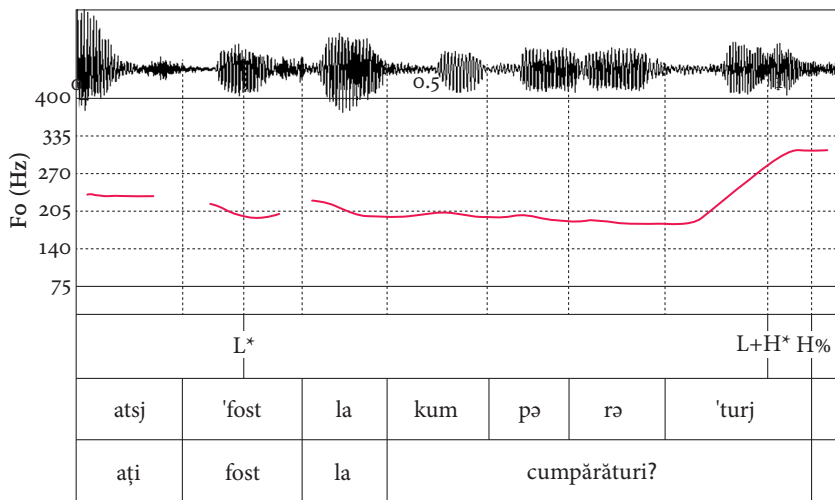


Figure 10. The B-pattern (EF, final syllable accented)

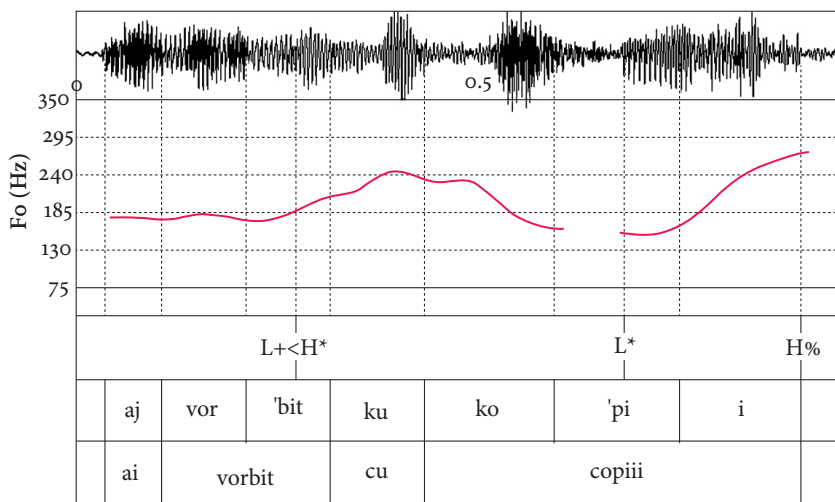


Figure 11. Pattern C (FE, last syllable unaccented)

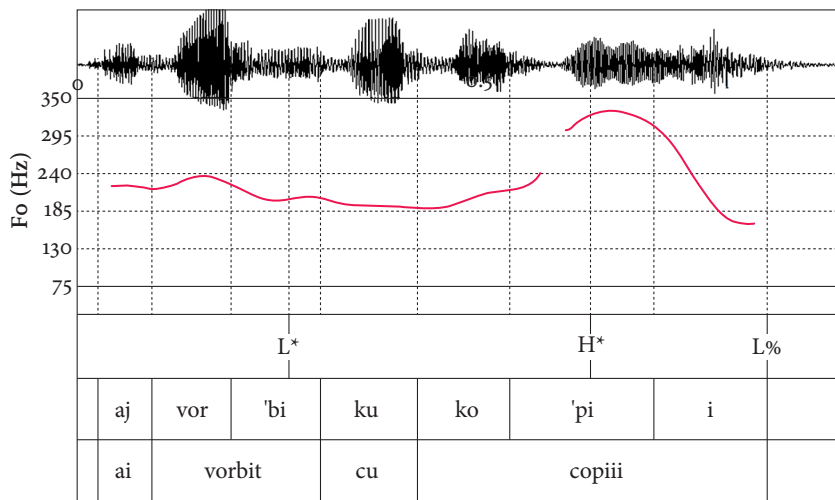


Figure 12. Pattern D (EF, last syllable unaccented)

The experimental results are given in Table 1. The patterns that are at odds with the predictions of my account are shaded:

Table 1. Experimental results

Informant:	Example (22):	Example (23):	Example (24):	Example (25):
1	A	B	E	D
2	A	B	C	D
3	A	B	D	D
4	A	A	C	D
5	A	A	C'	D
6	A	A'	C'	D
7	A	A	C	D
8	A	B	C	C
9	A	B	C	D
10	A	A	C	C
11	A	B	D	D
12	A	A	C	C
13	A	A	C	D

Although the predictions are not completely fulfilled, the results support the claim that the FF pattern is neutral. This is particularly clear for the pair in (22)–(23): here, all the unexpected results were found in the verum focus condition. We can conclude that some informants did not consider that the issue *whether p* was so much activated in the proposed context as to trigger the verum intonation (or, maybe, they didn't pay sufficient attention to the suggested context), so they chose the unmarked pattern. Note also that two informants (10 and 12) also used the FF pattern in (25), which shows that they were not sensitive to the contextual indications and consistently used the unmarked pattern. What remains to be explained are the unexpected results in the out-of-the-blue condition of the second pair (Example (24)). Note that all three informants also chose the EF (verum) pattern in (23). I suppose that they imagined a similar context for (24), where the issue *whether p* had previously occurred, although this was not indicated in the description of the context. All in all, we find most of the EF patterns in the verum condition (16 out of 18, if we do not count the isolated pattern E). The fact that the FF pattern extends beyond the out-of-the-blue condition into the verum condition can easily be explained as being due to the difficulty, for some speakers, to imagine the described context, and, anyway, supports the claim that this pattern is the unmarked one.

5. Other uses of the EF pattern with nuclear stress on V

Nuclear stress on the verb can also result from deaccenting the postverbal material. A context triggering deaccentuation of the Given is exemplified in (28). Note that both in the declarative and interrogative version of the sentence, the object is deaccented:

(28) [Context: Enescu's music is on the radio]

a. *Îmi 'PLA.ce E.'nes.cu.*

L+H* L- L%

"I like Enescu."

b. *Îți 'PLA.ce E.'nes.cu?*

L* L+H* L%

"Do you like Enescu?"

The EF pattern characterizing (28b) is shown in Figure 13.

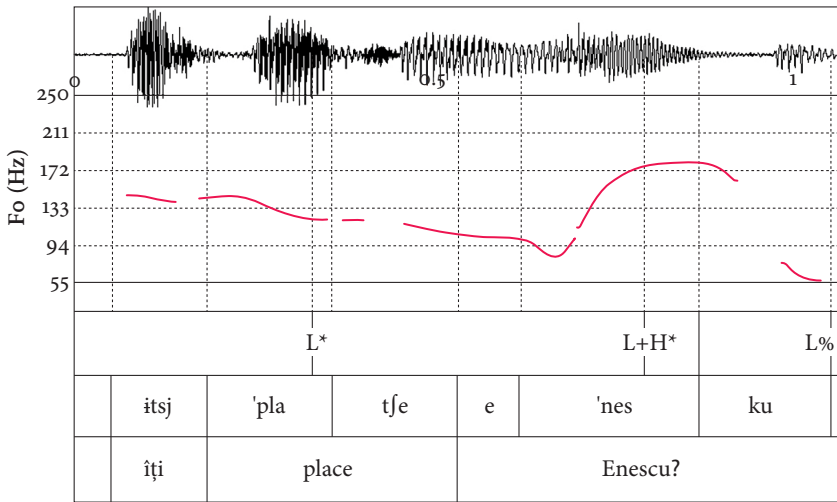


Figure 13.

The final HL% contour of EF correlated with a low tone on the verb also has a usage where it is not easily justifiable as reflecting verum focus or deaccenting: it appears in questions used to make proposals or requests – see (29), which is imagined as an out-of-the-blue polite request; its intonation is given in Figure 14.

- (29) *Ai putea să-mi împrumuți mașina?*
 would.2SG can SBJV-me.DAT lend.2SG car.the
 “Could you lend me your car?”

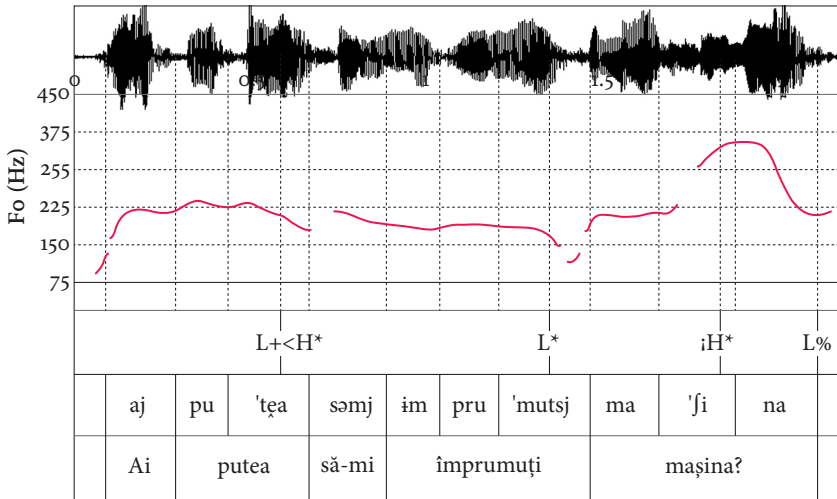


Figure 14.

There is however a difference with respect to the verum pattern: the L^* accent is not always on the main verb, but may occur on its subjunctive complement (see (30), which is imagined as an out-of-the-blue proposal, with the intonation given in Figure 15).

- (30) *Ai vrea să mergem la teatru?*
 would.2SG want SBJV go.1PL to theatre
 “Would you like to go to the theatre?”

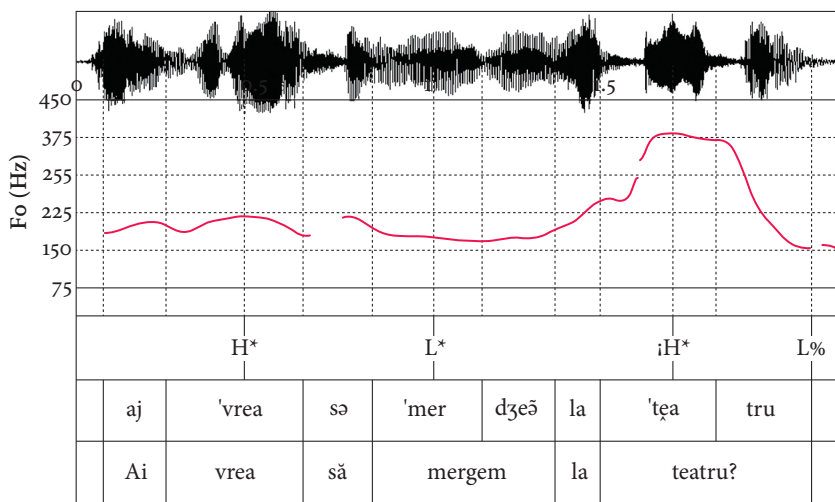


Figure 15.

The use of the FF pattern in such sentences would turn them into genuine information-seeking questions about a pre-existing intention of the hearer (instead of them being proposals or requests). As the EF pattern seems to be neutral here, we may have an explanation for the fact that this pattern has been considered *the* neutral pattern of polar questions in Romanian: the subjects may have considered questions used as requests or proposals rather than information-seeking questions.

A detailed investigation into the prosodic properties of questions with a directive import in Romanian has not yet been done. At this point, it is not clear to what extent their intonation really resembles the verum focus pattern and it is premature to seek an explanation of this similarity.

5. Conclusions

The prosodic characterization of polar questions in Romanian is first of all important from a descriptive point of view, given that intonation is the only systematic means of distinguishing polar questions from declaratives in this language (see Giurgea & Remberger 2012, 2014 for arguments against the view that polar questions may be indicated by word order). Secondly, the claim that the neutral intonation involves an early focus pattern with main prominence on the verb would be significant for the general theory of questions, suggesting a relation between a certain type of focus and polar question formation. The novel data presented in this chapter falsify this claim. I have argued that truly out-of-the-blue information-seeking polar questions have the nuclear stress at the end, as expected. The fact that in many cases we find a nuclear stress on the verb is the effect of verum focus: the speaker uses this prosody to signal that the issue *whether p* was already present among the questions under discussion.⁵

References

- Alboiu, Gabriela, Virginia Hill & Ioanna Sitaridou. 2015. Discourse driven V-to-C in early modern Romanian. *Natural Language & Linguistic Theory* 32. 4.
<https://doi.org/10.1007/s11049-014-9270-8> (online 2014), Nov; Nov 2015, 33: 4. 1057–1088.
- Avram, Andrei. 1973. Particularităţi ale intonaţiei interrogative în graiul din Muscel. *Fonetică şi dialectologie* 8. 43–64.
- Barbu, Ana Maria. 1999. Complexul verbal. *Studii şi cercetări lingvistice* 50(1). 39–84.
- Bianchi, Valentina, Giuliano Bocci & Silvio Cruschina. 2016. Focus fronting, unexpectedness, and evaluative implicatures. *Semantics and Pragmatics* 9(3). 1–54. <https://doi.org/10.3765/sp.9.3>

5. Further comparative research may lead to a finer-grained analysis of the situations I subsumed under ‘verum focus’. Thus, according to an anonymous reviewer, in German the counterparts of the Examples (23) and (25), which have a verum focus pattern in Romanian, do not have the nuclear stress on C/Spec,CP (the pattern characteristic of verum focus, see (9c)) but rather on the clause-final verb – for instance, the translation of (25) would sound *Hast du die Kinder GESPROCHEN?* ‘Have you SPOKEN to your children?’. Note that this is also a marked pattern (with internal arguments, the default stress is on the last preverbal argument; here, it should have been on *Kinder*). The reviewer explains this pattern by deaccenting of the complement due to givenness. But note that in the context at hand, both the complement *and the verb* are given (there was a previous issue of ‘speaking to the children’), so the stress on the verb cannot reflect the newness of the verb itself. This means that, like the stress on the auxiliary used in standard instances of verum focus, this stress reflects focus on a clausal functional category. Eva Remberger (p.c.) reports that main stress on the auxiliary is also possible in this example, but involves more insistence.

- Bianchi, Valentina & Silvio Cruschina. 2016. The derivation and interpretation of polar questions with a fronted focus. *Lingua* 170. 47–68. <https://doi.org/10.1016/j.lingua.2015.10.010>
- Boersma, Paul & David Weenink. 2012. *Praat – Doing phonetics by computer* [Computer program]. <http://www.praat.org/> (13 June, 2016)
- Büring, Daniel. 2006. Intonation und Informationsstruktur. In Hardarik Blühdorn, Eva Breindl & Ulrich H. Waßner (eds.), *Text-Verstehen. Grammatik und darüber hinaus*, 144–163. Berlin: de Gruyter.
- Dascălu-Jinga, Laurenția. 2001. *Melodia vorbirii în limba română*. Bucharest: Univers Enciclopedic.
- Dascălu, Laurenția. 1979a. Teste de percepție asupra intonației în limba română. *Studii și cercetări lingvistice* 30(2). 125–139.
- Dascălu, Laurenția. 1979b. On the intonation of questions in Romanian: The rising pattern. *Revue roumaine de linguistique* 24(1). 35–44.
- Dik, Simon, Sie Ing Djiang, Maria Hoffman, Jan de Jong, Harry Stroomer & Lourens de Vries. 1981. On the typology of focus phenomena. In Teun Hoekstra, Harry van der Hulst & Michael Moortgat (eds.), *Perspectives on functional grammar*, 41–74. Dordrecht: Foris.
- Dobrovie-Sorin, Carmen. 1994. *The syntax of Romanian*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110886597>
- Dobrovie-Sorin, Carmen & Charlotte Galves. 2000. Proclisis, enclisis and head-to-head merge. *Bucharest Working Papers in Linguistics* 2(1). 35–51.
- GALR = Academia Română, Institutul de Lingvistică, Iorgu Jordan – Al. Rosetti. Valeria Guțu Romalo (ed.), *Gramatica limbii române*. Bucharest: Editura Academiei Române, 2008 (2 volumes).
- Giurgea, Ion. 2016. On the interpretation of focus fronting in Romanian. *Bucharest Working Papers in Linguistics* 18(2). 37–61.
- Giurgea, Ion & Eva-Maria Remberger. 2012. Verum focus and polar questions. *Bucharest Working Papers in Linguistics* 14(2). 21–40.
- Giurgea, Ion & Eva-Maria Remberger. 2014. Polarity fronting in Romanian and Sardinian. In Marie-Hélène Côté & Eric Mathieu (eds.), *Variation within and across Romance languages. Selected papers from the 41st Linguistic Symposium on Romance Languages (LSRL), Ottawa, 5–7 May 2011*, 173–197. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.333.13giu>
- Grice, Martine, Robert Ladd & Amalia Arvaniti. 2000. On the place of phrase accents in intonational phonology. *Phonology* 17. 143–185. <https://doi.org/10.1017/S0952675700003924>
- Gussenhoven, Carlos. 1984. *On the grammar and semantics of sentence accents*. Berlin: Walter de Gruyter. <https://doi.org/10.1515/9783110859263>
- Gutzmann, Daniel & Elena Castroviejo Miró. 2011. The dimensions of verum. In Olivier Bonami & Patricia Cabredo Hofherr (eds.), *Empirical issues in syntax and semantics* 8, 143–165. <http://www.cssp.cnrs.fr/eiss8> (13 June, 2016)
- Gutzmann, Daniel, Katharina Hartmann & Lisa Matthewson. Manuscript. Verum focus is verum, not focus. Cross-linguistic evidence. 2017. <https://www.danielgutzmann.com/work/verum-focus-is-verum-not-focus> (13 June, 2016)
- Halliday, Michael. A. K. 1967. *Intonation and grammar in British English*. The Hague: Mouton. <https://doi.org/10.1515/9783111357447>
- Höhle, Tilman. 1988. Vorwort und Nachwort zu Verumfokus. *Sprache und Pragmatik* 5. 1–7.
- Höhle, Tilman. 1992. Über Verumfokus im Deutschen. In Joachim Jacobs (ed.), *Informationsstruktur und Grammatik*, 112–141. Opladen: Westdeutscher Verlag. https://doi.org/10.1007/978-3-663-12176-3_5

- Holmberg, Anders. 2013. The syntax of answers to polar questions in English and Swedish. *Lingua* 128. 31–50. <https://doi.org/10.1016/j.lingua.2012.10.018>
- Jitcă, Doina, Vasile Apopei, Otilia Păduraru & Samuil Marușca. 2015. Transcription of Romanian intonation. In Sonia Frota & Pilar Prieto (eds.), *Intonation in Romance*, 284–316. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199685332.003.0008>
- Ladd, Robert. 2008. *Intonational phonology* (2nd edn.). Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511808814>
- Lohnstein, Horst. 2012. Verumfokus – Satzmodus – Wahrheit. In Horst Lohnstein & Hardarik Blühdorn (eds.), *Wahrheit – Fokus – Negation*, 31–67. Hamburg: Buske Verlag.
- Lohnstein, Horst. 2016. Verum focus. In Caroline Féry & Shinichiro Ishihara (eds.), *The Oxford handbook of information structure*, 290–313. Oxford: Oxford University Press.
- Repp, Sophie. 2013. Common ground management: Modal particles, illocutionary negation and verum. In Daniel Gutzmann & Hans-Martin Gärtner (eds.), *Beyond expressives: Explorations in use-conditional meaning*, 231–274. Leiden: Brill. https://doi.org/10.1163/9789004183988_008
- Romero, Maribel. 2005. Two approaches to biased *yes/no* questions. In John Alderete, Chunghye Han & Alexei Kochetov (eds.), *Proceedings of the 24th West Coast Conference on Formal Linguistics (WCCFL)*, 352–360. Somerville, MA: Cascadilla Proceedings Project.
- Romero, Maribel & Chunghye Han. 2004. On Negative y/n-Questions. *Linguistics and Philosophy* 27. 609–658. <https://doi.org/10.1023/B:LING.0000033850.15705.94>
- Watters, John Robert. 1979. Focus in Aghem. In Larry Hyman (ed.), *Aghem grammatical structure*, 137–197. Los Angeles: University of Southern California.

The downward grammaticalisation of irrealis subordinators in Romanian, Salentino and southern Calabrese

Kim A. Groothuis
Ghent University

Romanian *să*, Salentino *cu* and southern Calabrese *mu/ma/mi* are irrealis subordinators, which also replace the infinitive in many of its uses. The present chapter investigates the diachrony of these elements. It is shown that they are all instances of downward grammaticalisation, deriving from high C-elements which move down to Fin⁰, and in the case of *cu* and *mu*, also further down to T- and *v*-related positions. QUOMODO is proposed as an etymon for both *cu* and *mu*, rather than the traditionally assumed QUOD and MODO (UT). QUOMODO, originally replacing UT, follows Haspelmath's (1989) grammaticalisation path for infinitives when developing into *cu* and *mu*. The development of SI into *să* is similar: it develops from a conditional complementiser to general irrealis subordinator.

Keywords: complementisers, Romanian, southern Calabrese, Salentino, downward grammaticalisation

1. Introduction

On a par with languages from the Balkan Sprachbund (Joseph 1983), a number of Romance varieties are characterised by a limited use of infinitival complementation. Instead, they make use of finite complements introduced by a dedicated 'subjunctive' or 'irrealis'¹ subordinator: Romanian *să* (1), Salentino *cu* (2), and *mu* (with its diatopic variants *u*, (*m*)*i*, and *ma*) in southern Calabrese and northeastern Sicilian

1. Irrealis modality is here to be understood in opposition with both realis (asserted by the speaker or matrix subject) and factive complements (which are presupposed), cf. Hooper & Thompson (1973). Irrealis complements refer to unrealised events which are neither asserted nor presupposed.

(3) (Sorrento 1951; Rohlfs 1969; Calabrese 1992; Dobrovie-Sorin 1994; Ledgeway 1998, 2013, 2016a; Alboiu & Motapanyane 2000; Damonte 2011; De Angelis 2017). These clauses are often considered subjunctives, but only Romanian (only in the 3rd person and with the verb *a fi* “to be”) and some Salentino dialects have verb forms morphologically marked as such (Bertocci & Damonte 2007).

- (1) *Vreau să merg la petrecere.*
 want.1SG să go.1SG to party
 “I want to go to the party.” (Ro.)
- (2) *Lu Karlu ole ku bbene krai.*
 the Karlu want.3SG CU come.3SG tomorrow
 “Karlu wants to come tomorrow.”
 (Sal. (Campi Salentina), Calabrese 1993: 28)
- (3) *Vogghiu mi veni.*
 want.1SG MU come.3SG
 “I want him to come.” (Sic. (Messina), Leone 1995: 68)

The present chapter focuses on diachrony of these three subordinators and aims to show that they present a case of downward (secondary) grammaticalisation as they derive from C-elements which moved down to Fin⁰, and in the case of *cu* and *mu*, further down the clausal spine.

2. Structural position of *mu*, *cu*, and *să*

Before discussing the diachrony of *mu*, *cu* and *să*, it is useful to look at their current structural position. They have been analysed variously as modal particles (Rivero 1994; Paoli 2003; Damonte 2011) or complementisers (Dobrovie-Sorin 1994: 93; Nicolae 2015: 133; Hill & Alboiu 2016: 240–245), because they show a dual character, heading complement clauses on the one hand (as in the examples 1–3) but also co-occurring with C-elements (except for *cu*) on the other:

- (4) *Vreau ca ION să meargă la petrecere.*
 want.1SG that Ion să go.SBJV.3SG to party
 “I want ION to go to the party.” (Ro.)
- (5) *Chinnommu cadì mai malatu!*
 that=NEG=MU fall.3SG ever ill
 “May s/he never fall ill.” (SCal., Roberts & Roussou 2003: 91)

In Romanian, *ca* occurs when a constituent is moved to the left periphery (Gheorghe 2013: 470).

Here, I adopt the view that *cu* and *mu* lexicalise different positions along the clausal spine, namely in the *v*P, TP or CP, depending on the matrix verb that selects the complement (Ledgeway 2012, 2015: 157; Taylor 2014; Squillaci 2016; Groothuis 2020: chap. 2). This same distribution is also found with Romance infinitival complementisers AD and DE (Ledgeway 2016a: 1014–1015), and Serbo-Croatian *da* (Todorović & Wurmbrand 2016). Lower functional verbs within the TP domain select smaller complements; higher functional verbs will select complements with more structure. According to this view, *mu* and *cu* lexicalise a head in the *v*-domain when subcategorised by root modal and lower aspectual predicates, whereas they encode a T-related head when subcategorised by epistemic/alethic modals, temporal and higher aspectual predicates, and that, finally *mu* and *cu* lexicalise a C-related head when subcategorised by lexical control predicates:

- (6) a. *Speramu* [_{CP} *armenu u focu nomm'u ddumanu stasira*].
 wish.1PL at.least the fire NEG=MU=it light.3PL tonight
 “Let’s hope that they don’t light the bonfire at least tonight.”
- b. *Cercu sempri* [_{TP} *nommi fumu*].
 try.1SG always NEG=MU smoke.1SG
 “I always try not to smoke.”
- c. *Finiscinu* [_{vP} *m'u mbivinu*].
 it=finish.3PL MU=it drink.3PL
 “They finish drinking it.”

(SCal. (Bova Marina), adapted from Squillaci 2016: 157)

Evidence for this can be gathered by applying a series of tests (including transparency tests, movement and pronominalisation of the complement) on three types of verbs: aspectuals, modals and lexical control verbs (see Groothuis 2020: Chapter 2 for discussion). On the basis of these tests, I have concluded that Salentino *cu* and southern Calabrese *mu* do indeed occupy various positions along the clausal spine; Romanian *să*, on the other hand, behaves differently as it systematically lexicalises Fin (Stan 2007; Gheorghe 2013; Nicolae 2016; Hill & Alboiu 2016: 240–145).

3. A unified etymology for Salentino *cu* and southern Calabrese *mu*

3.1 Previous etymologies of *cu* and *mu*

The traditionally accepted etymologies for *cu* and *mu* are, respectively, the Latin complementiser QUOD (Rohlf 1969; Loporcaro 1997: 347; Mancarella 1998), and Latin adverb MODO “now, presently”, which in turn derives from the ablative of MODUS “way, manner” (Sorrento 1951; Rohlf 1969: 192; Ledgeway 1998; Roberts

& Roussou 2003: 88; De Angelis 2013, 2017). With regard to *MODO* as the source of *mu*, there are two main hypotheses. The first one is that *MODO* occurred in paratactic constructions, such as the following:

- (7) *Volo et modo venio.*
 want.1SG and now come.1SG
 “I want and now I come” > “I want to come.” (Lat., Ledgeway 1998: 48)

From this paratactic use, the subordinator *mu* would have originated. It has been noted, however, that there is no evidence for such a use of *MODO* (De Angelis 2017: 142). It is also not clear why this paratactic use should be limited to irrealis complements.

The second hypothesis is that the combination of *MODO UT*, expressing counterfactuality (8), grammaticalised into *mu* (Roberts & Roussou 2003: 93–95).

- (8) *Modo ut sciam.*
 modo that know.SBJV.1SG
 “If only I knew.” (Lat., Roberts & Roussou 2003: 94)

Nevertheless, this etymology is not without problems either: *ut* has been lost in the transition from Latin to Romance, whereas *mu* only grammaticalised between the 5th and the 11th centuries (Roberts & Roussou 2003: 97 fn.8). It is therefore very implausible that it would only survive in this combination in only the Greek speaking areas of Calabria. Furthermore, counterfactuality in southern Calabrese is not expressed with *mu* but with *si* “if” (Chillà & Citraro 2012: 118):

- (9) *S’u sapiva ...*
 if=it know.IPFV.1SG
 “If only I knew.” (SCal. (Bova))

This is unexpected if the counterfactual value is the context in which *MODO UT* has grammaticalised into *mu*.

3.2 QUOMODO

A different approach, followed here, is suggested by Bertoni (1905, *apud* De Angelis 2016: 77 n.8) and adopted by Ledgeway (2016b), according to which both *cu* and *mu* derive from *QUOMODO* “how”. *Quomodo* is also given as an etymon for *mu* by Meyer-Lübke (1899, III: 516) and Scerbo (*apud* Sorrento 1951). There are several reasons why *QUOMODO* is a more probable etymon than *QUOD* or *MODO (UT)*.

First, unlike *modo*, *quomodo* is already used as a complementiser in late Latin in declarative contexts, and it could replace *ut* “so that” as a purposive complementiser, which does not survive into Romance, see (10):

- (10) *Columbaria singula esse oportet, ut os
nest.NOM singular.NOM be.INF be-necessary.3SG so mouth.ACC
habeat (columba), quo modo inire et exire possit
have.SBJV.3SG dove.NOM how enter.INF and exit.INF can.SBJV.3SG*
“Every nest needs to be [like this], so that it has a mouth so that (the dove) can enter and exit.” (Lat., Varro *Rust.* 3, 7, 4 *apud* Hofmann & Szantyr 1972: 650)

Both *cu* and *mu* can head a purposive adjunct clause, particularly when the matrix verb indicates motion.

Second, cross-linguistically, *how* has been attested elsewhere to develop into a complementiser (Willis 2007; van Gelderen 2015). In contrast, adverbs such as *modo* “now” develop more rarely into complementisers. From a cross-linguistic perspective, *quomodo* is thus a more probable etymon. Moreover, the use of the reflexes of QUOMODO as a purposive complementiser is attested in several old Romance varieties (Meyer-Lübke 1899, III: 641), as in (11):

- (11) *Io m'aggio posto in core a Dio servire, com' io potesse
I to.me=have.1SG put in heart to God serve.INF how I can.SBJV.IPFV.1SG
gire in paradiso.
go.INF to paradise*
“I resolved to serve God, so that I could go to paradise.”
(OIt., Giacomo di Lentini, *apud* Rohlfs 1969: 181)

This shows that QUOMODO has been used in a similar context as the ones in which *cu* and *mu* appear in southern Calabrese and Salentino, namely heading purposive clauses.

Finally, in order to allow Salentino *cu* to derive from QUOD, it is necessary to make extra assumptions for early Salentino, which is characterised by a triple complementation system, featuring *cu*, *che*, and *ca* (cf. Sgrilli 1984: 160–165; Ledgeway 2005: 367–370). *Che* (and *que* in other Romance varieties) is generally assumed to be the result of a merger of the interrogative pronoun QUID and relative pronoun/complementiser QUOD (cf. Rohlfs 1969: 188, among others), but if QUOD gave *cu*, we have to assume this merger did not take place in Salentino. Instead, if we assume QUOMODO > *cu*, Salentino *che* can have the same etymology as other cognates in Romance.

I conclude therefore that QUOMODO is a more probable etymon for both *cu* and *mu* than the traditionally assumed QUOD and MODO (UT). The next subsection will trace the development of QUOMODO in more detail.

3.3 The development of QUOMODO > *cu/mu*

Quomodo was originally a compound from the interrogative *quis* “which” and the noun *modus* “way, mode”, both in the ablative case, meaning “in which way”. Being a *wh*-element, *quomodo* is a phrasal element that in the course of the derivation moves to Spec,CP or, within the split CP, to Spec,FocusP (Rizzi 1997). On the other hand, *cu* and *mu* are functional heads that can occupy different positions along the clausal spine depending on the matrix verb that selects them. How can we account for this change?

Ut “(so) that” (and its negative counterpart *ut non/ne*) was the main irrealis and purposive complementiser in classical Latin, which was however lost in the transition to Romance as it had become weak both on phonologically and semantically (Herman 1963: 53; Hofmann & Szantyr 1972: 632, 646; Vincent 1988: 68). In late Latin, other C-related elements start taking over functions of *ut*, including *quod* “that” and *quomodo* “how”. Given the shared meaning “how” between *quomodo* and *ut*, it is not unexpected that *quomodo* by analogy takes over the other functions of *ut*, including the function as final and irrealis complementiser.

The reanalysis of the *wh*-phrase *quomodo* as a C-head can be explained by two economy principles: the Head over Phrase principle, according to which it is more economical for language-acquiring children to posit a head than a phrase (Van Gelderen 2009), and the Merge over Move Principle (Roberts & Roussou 2002; cf. also Van Gelderen’s (2009) Late Merge Principle), which states that it is less costly to merge an element in a higher position directly than to move it there from a lower position within the tree. Children acquiring Latin see *quomodo* as synonym of *ut* when the latter is a *wh*-element; by analogy, they extend *quomodo* also to the other uses of *ut*. In the case of irrealis complementiser, rather than being a moved *wh*-element, it is more economical to posit a head which is directly merged into the C-domain with an [irrealis] feature.

This reanalysis from phrase to head leads us to expect it to lose some of its morphophonological structure (viz. > *quo* or > *mo(do)*). That both the first and the second part of the *wh*-word were conserved can be explained by the compound nature of QUOMODO. This compound must have been transparent, as both *quis* and *modus* remain frequent throughout the history of Latin, and the meaning of the compound is compositional. Indeed, *quomodo* could appear in *tnesis* (cf. Lewis, *Short Oxford Latin Dictionary* ‘quomodo’):

- (12) *quo tu me modo voles esse*
 which.ABL YOU.NOM me.ACC way.ABL want.SBJV.3SG be.INF
 “As you want me to be” (Lat., Plaut. *Cist.* 1, 1, 48)

We hypothesise thus that both parts survived, *quo*- in Salentino, yielding *cu*, and *-mo(do)* in southern Calabrese, giving *mu*.

The stress of *quōmōdō* was *quómodo* rather than *quomódo*. This is unproblematic for deriving *cu*, which supposedly derives from *quo* after *-modo* has been dropped. Assuming that *mu* derives from the second is not problematic either. Given the transparent nature of this compound, of which *modo* is the head, we can assume *modo* still bears some stress. Furthermore, when appearing in *tnesis* as in the example above, *modo* must have borne stress. Being the head of the compound, it is not surprising that in the Calabrese varieties *mo(do)* has been retained rather than *quo-*.

QUOMODO is a more marked form to use as irrealis complementiser than the general complementiser QUOD/QUE, which can be used with any type of complement, whereas final QUOMODO introduces irrealis (purpose) clauses. In fact, in many other southern Italo-Romance varieties, QUOD/QUE introduces irrealis complements (in opposition with realis QUIA > *ca*). Why do the extreme southern Italian varieties opt for QUOMODO instead?

The choice for the more marked option must be linked to the influence of Greek. As in many Balkan languages, the infinitive in Greek has been lost and replaced by a finite clause headed by the final complementiser (*hi*)*na* (see Joseph 1983 Chapter 3, among many others). Typologically, infinitives tend to derive from purposive constructions, which in turn often derive from allative constructions, cf. English *to* and German *zu* (Haspelmath 1989), as well as the Romance *a/à* introducing irrealis infinitival clauses. Infinitives share with purposive clauses their irrealis, unrealised character (Stowell 1982; Haspelmath 1989). It is therefore not surprising that when the infinitive disappeared, it was substituted by a final clause.

The intense language contact and widespread bilingualism has led to the structural borrowing of this phenomenon from Greek in the extreme southern Romance varieties (cf. Rohlfs 1969: 190, among many others). *Quomodo* is a perfect candidate to mirror (*hi*)*na*: apart from being used as an irrealis complementiser, it also has the purposive meaning that characterises (*hi*)*na*. *Quod/que*, on the other hand, is an unmarked clause linker.

Haspelmath (1989) argues for the following grammaticalisation path of infinitives from purposive clauses:

- (13) Purposive > irrealis directive modality (manipulative and volitional verbs) > irrealis-potential (modals and evaluative verbs) > irrealis-(non)factive (thinking and verbs of utterance), factive (cognition and evaluative predicates).

(Haspelmath 1989: 298–299)

On the basis of this hierarchy, we assume that the infinitive was first replaced by *quomodo* in the purposive contexts, after which it was extended to irrealis complements in general. This is indeed confirmed by Ledgeway's (2013: 200) results, who finds that purposive contexts after movement verbs such as *come* and *go* are

replaced almost everywhere by finite complementation in Calabrese and Salentino; other irrealis-potential complements still (optionally) take the infinitive.

The grammaticalisation of *cu* and *mu* seems to follow the cross-linguistically frequent grammaticalisation path of infinitives. This means that *cu* and *mu* acquire the possibility to be merged in lower positions, not only in a C-related head (arguably Fin), but also in the T- and *v*-domain. Since *cu* and *mu* occupy various positions along clausal spine, this is a case of downward grammaticalisation.²

4. The development of Romanian *să*

Să (in old Romanian also written *se*) is standardly argued to derive from the Latin conditional complementiser *si* ‘if’ (Herman 1963: 175; Jordan 2009: 25; Nicolae 2015; Zafiu et al. 2016). The semantic change that *să* has undergone from conditional complementiser to subordinating subjunctive particle is however unclear (Zafiu et al. 2016: 15). Frîncu (1969) and Hill (2013) argue that *si* became an irrealis complementiser and later a mood marker. On the other hand, Zafiu et al. (2016: 15) argue that it is more likely that *să* also had a purposive meaning in old Romanian, derived from the Latin adverbial use of *se/si(c)*. Here, we propose that this latter homophony between the conditional complementiser and the adverbial *se/si* from < *si(c)* that is crucial for the development of *să*. The adverbial *se/si(c)* was used to introduce purposive clauses, as does modern-day *să*. Second, in some varieties, *să* is homophonous to, or can be replaced by, *și* ‘and, too’, which derives from Lat. *sic* (Nedelcu et al. 2016: 17). We therefore assume that the element *se*, resulting from the homophony between Latin *si* and *si(c)*, has grammaticalised further into the subjunctive subordinating particle *să*. This element was marked for [irrealis] and could also introduce purpose clauses, as it still does today, and it mirrors the irrealis/purposive nature of infinitives (Haspelmath 1989).

Se/să in old and modern Romanian generally replaces the infinitive in complementation, leading to a change that is attested in most of the Balkan languages (Sandfeld 1930; Joseph 1983). Indeed, Romanian employs finite complements introduced by *să* where most other Romance languages would use an infinitive. However, the infinitive is not a verb form unknown to modern Romanian, and is still regularly used, even if to a very limited degree, with modal verbs as *a putea* ‘to be able to’ and in temporal and modal periphrases (Zafiu 2013).

The grammaticalisation of *să* differs from that of *cu* and *mu* because old Romanian *se* is already a C-related head. Rizzi (2001) argues Italian *se* ‘if’ is located

2. cf. also Andriani, Groothuis & Silvestri (2020), who discuss the development of *cu* and *mu* in the wider context of grammaticalisation pathways in Italo-Romance.

in IntP. We can assume the same for Latin/Romance *si*. We have argued above that Romanian *să* on the other hand is in Fin. Therefore, the modern use of *se/să* as a subjunctive/irrealis complementiser is lower than conditional *să*, which has been lost in modern Romanian. This development from conditional to irrealis complementiser is thus a case of downward grammaticalisation.

5. Downward grammaticalisation

Grammaticalisation “involves the creation of new functional material, either through the reanalysis of existing functional material or through the reanalysis of lexical material” (Roberts & Roussou 2003: 2). This type of diachronic change is often characterised by semantic bleaching and phonological reduction of the element undergoing grammaticalisation.

In the cases discussed here, we are dealing with a specific subtype of grammaticalisation, as the elements undergoing grammaticalisation (*QUOMODO* > *cu*, *mu* and *si* > *să*) are already grammatical elements belonging to the C-domain in Latin. It might therefore be more correct to speak about ‘secondary grammaticalisation’, although there is discussion whether this processes really differ from ‘standard’ grammaticalisation (Brinton & Traugott 2005: 53; Breban 2014).

In the generative approach to grammaticalisation as developed by Roberts & Roussou (2003), grammaticalisation is considered as an ‘upward’ process: the lexical (or functional) material is reanalysed as merged directly into a higher position within the clause instead of moving there. Apparent ‘downward’ changes involve loss of movement; these are not instances of grammaticalisation and have the following properties which distinguish them from actual cases of grammaticalisation (Roberts & Roussou 2003: 208):

- i. Apply to all members of category Y;
- ii. Do not change the category of Y;
- iii. Involve no semantic or phonological change to Y-roots;
- iv. Cannot be cyclic.

Taking the diachrony of *cu*, *mu* and *să* in consideration, we see that these properties are not attested. The proposed grammaticalisation path does not apply to all C-elements (cf. realis complementisers *ca*, *că*, *chi* which are still in Force), but only to the subset of irrealis complementisers in these varieties. Their etyma do change in category; particularly in the case of *QUOMODO*, as *mu* and *cu* can occupy positions in *v*- and T-domain as well. Morphophonological erosion is clearly attested in the development from *QUOMODO* to *cu* and *mu*. Rather than semantic bleaching in the strict sense, the grammatical elements take on new (broader) functions (from

final or conditional complementisers to general irrealis subordinators). Also point (iv) seems to be disconfirmed by the further development of *cu* and *mu* which can appear in lower positions in the clause as well, not just within the C-domain. The elements investigated are thus proper instances of grammaticalisation and cannot be dismissed as cases of loss of movement.

Roberts & Roussou (2003: 97) notice the problematic status of *mu* (and Greek *na*); they argue that the grammaticalisation of Greek *na* and Calabrese *mu* are not cases of downward grammaticalisation, because the modality features are no longer realised in the I domain (on verbs) but in C domain (*na* is in Fin and can optionally move to C). In this sense, the features have moved upwards and the direction of the change is still upwards. Indeed, the Calabrian varieties do not present a separate paradigm for subjunctive forms; after *mu*, regular indicative verb forms are employed.

A similar situation is found in Romanian, where only the 3rd person subjunctives present a distinct morphological verb form, and can under certain circumstances be realised without *să*. However, apart from these marked forms, since originally *si/se* as conditional marker is merged in a higher position than subjunctive *să* (which could be in Fin or in the IP-domain), there is still a grammaticalised element ‘moving down’, in the sense that it is reanalysed as lexicalizing a lower head in the tree.

Roberts & Roussou’s solution cannot straightforwardly be extended to the case of Salentino *cu* either, as in many Salentino varieties (at least some) verbs still present distinctive subjunctive forms (Bertocci & Damonte 2007). In these varieties, mood is still marked on the T head, as well as on the *cu* or *mu*. Furthermore, as discussed in § 2, *cu* can lexicalise different positions along the clausal spine; when lexicalizing a T- or *v*-related position, the modal features are lexicalised in these domains and hence lower than the CP.

These can therefore be considered cases of downward grammaticalisation: $si(c) > s\check{a}$, $QUO(MODO) > cu$ and $(QUO)MODO > mu$ are reanalysed to a lower position within the C-domain, and, in the case of *mu* and *cu*, to T- or *v*-related positions as well. The diachrony of all three subordinators constitutes therefore an exception to the generalisation made by Roberts & Roussou (2003). Examples of downward grammaticalisation have been described before. Specifically within the C-domain, Munaro (2016) describes cyclic downward grammaticalisation of irrealis *che* in some Italo-Romance varieties. The question arises why specifically *irrealis* complementisers seem to grammaticalise downwards; this might be related to the fact that functional complements, which are smaller in Romance, often pattern together with irrealis complements, being realised by infinitives or subjunctives. Another reason might have to do with the fact that these are all instances of grammatical elements assuming new grammatical functions. The upward movement might be obligatory only when *lexical* elements grammaticalise. Further research is needed at this point (see Groothuis 2020: Chapter 3 for discussion).

6. Conclusions

From the preceding discussion it emerges that although they have a different origin, the irrealis subordinators found in Romanian, Salentino, and Calabrese show a similar development as they are all instances of downward grammaticalisation. For Salentino *cu* and Calabrese *mu* a new etymon has been proposed: Latin QUOMODO ‘how’.

The three cases discussed above are all instances of downward grammaticalisation, because both QUOMODO and SI(C) are originally located high in the C-domain. During their grammaticalisation to irrealis subordinators, they move down to at least Fin, and even further in the case of *cu* and *mu*. As these can be used with more functional verbs, they head smaller complements and so ‘move down’ along the clausal spine. I have argued that these are instances of grammaticalisation and not just loss of movement.

Acknowledgements

I wish to thank Adam Ledgeway for his feedback and for encouraging me to pursue this line of research. I would also like to thank Maria-Olimpia Squillaci and Alexandru Nicolae for their comments and discussion of the data from Calabrese and Romanian respectively. Finally, I wish to thank the audiences at the University of Bucharest, Trinity Hall, and UCL, where previous versions of this paper have been presented.

References

- Alboiu, Gabriela & Virginia Motapanyane. 2000. The generative approach to Romanian grammar: An overview. In Virginia Motapanyane (ed.), *Comparative Studies in Romanian Syntax*, 1–48. Amsterdam: Elsevier.
- Andriani, Luigi, Kim A. Groothuis & Giuseppina Silvestri. 2020. Pathways of grammaticalisation in Italo-Romance. (Ed.) Adam N. Ledgeway & Ian G. Roberts. *Probus* 32(2). 327–366. <https://doi.org/10.1515/prbs-2020-0004>
- Bertocci, Davide & Federico Damonte. 2007. Distribuzione e morfologia dei congiuntivi in alcune varietà salentine. In Federico Damonte & Jacopo Garzonio (eds.), *Studi sui dialetti della Puglia*, 3–28. Padua: Unipress.
- Bertoni, Giulio. 1905. Intorno alle questioni sulla lingua della lirica italiana. *Studi Medievali* 1. 580–593.
- Breban, Tine. 2014. What is secondary grammaticalization? Trying to see the wood for the trees in a confusion of interpretations. *Folia Linguistica* 48(2). 469–502. <https://doi.org/10.1515/flin.2014.016>
- Brinton, Laurel J. & Elizabeth Closs Traugott. 2005. *Lexicalization and language change*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511615962>

- Calabrese, Andrea. 1992. The lack of infinitival clauses in Salentino: A synchronic analysis. In Christiane Laeuffer & Terrell A. Morgan (eds.), *Theoretical analyses in Romance linguistics. Selected papers from the 19th Linguistic Symposium on Romance Languages (LSRL XIX). The Ohio State University, 21–23 April 1989*, 267–294. Amsterdam: John Benjamins.
<https://doi.org/10.1075/cilt.74.19cal>
- Calabrese, Andrea. 1993. The sentential complementation of Salentino: a study of a language without infinitival clauses. In Adriana Belletti (ed.), *Syntactic Theory and the Dialects of Italy*, 28–98. Turin: Rosenberg & Sellier.
- Chillà, Leonida & Cinzia Citraro. 2012. Impopolarità dell'infinito nelle varietà calabresi meridionali estreme: la selezione dei complementatori ei contesti modali. *Rivista italiana di linguistica e di dialettologia* 14(14). 113–144.
- Damonte, Federico. 2011. Matching moods: Mood concord between CP and IP in Salentino and southern Calabrian subjunctive complements. In Paola Benincà & Nicola Munaro (eds.), *Mapping the left periphery: The cartography of syntactic structures*, 228–256. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199740376.003.0008>
- De Angelis, Alessandro. 2013. *Strategie di complementazione frasale nell'estremo meridione italiano*. Messina: SGB Edizioni.
- De Angelis, Alessandro. 2016. Origini formali e funzionali della particella (*m*)i, (*m*), *ma* nell'area messinese e calabrese centro-meridionale. In Patrizia Del Puente (ed.), *Dialetti: per parlare e parlarne. Atti del IV Convegno Internazionale Di Dialettologia. progetto a.l.ba. (Potenza, Castelmezzano, Lagopesole, 6–8 novembre 2014)*, 75–94. Potenza: Il Segno Arti Grafiche.
- De Angelis, Alessandro. 2017. Between Greek and Romance: Competing complementation systems in southern Italy. In Piera Molinelli (ed.), *Language and Identity in Multilingual Mediterranean Settings*, 135–156. Berlin, Boston: De Gruyter Mouton.
<https://doi.org/10.1515/9783110554274-008>
- Dobrovie-Sorin, Carmen. 1994. *The syntax of Romanian*. Berlin: Mouton de Gruyter.
<https://doi.org/10.1515/9783110886597>
- Frîncu, Constantin. 1969. Cu privire la “Uniunea Lingvistică Balcanică”. Înclocuirea infinitivului prin construcții personale în limba română veche. *Anuar de Lingvistică și Istorie Literară* 20. 69–116.
- Gelderen, Elly van. 2009. Renewal in the left periphery: Economy and the complementizer layer. *Transactions of the Philological Society* 107(2). 131–195.
<https://doi.org/10.1111/j.1467-968X.2009.01216.x>
- Gelderen, Elly van. 2015. The particle *how*. In Josef Bayer, Roland Hinterhölzl & Andreas Trotzke (eds.), *Discourse-oriented syntax*, 159–174. Amsterdam: John Benjamins.
<https://doi.org/10.1075/la.226.07gel>
- Gheorghe, Mihaela. 2013. The structure of complex clauses. Subordination. argument clauses. In Gabriela Pană Dindelegan (ed.), *The grammar of Romanian*, 466–473. Oxford: Oxford University Press.
- Groothuis, Kim A. 2020. Reflexes of finiteness in Romance. Unpublished doctoral dissertation, University of Cambridge.
- Haspelmath, Martin. 1989. From purposive to infinitive – A universal path of grammaticization. *Folia Linguistica Historica* 23. 287–310. <https://doi.org/10.1515/flih.1989.10.1-2.287>
- Herman, József. 1963. *La formation du système roman des conjonctions de subordination*. Berlin: Akademie Verlag.
- Hill, Virginia. 2013. The emergence of the Romanian subjunctive. *The Linguistic Review* 30(4). 1–37. <https://doi.org/10.1515/tlr-2013-0016>

- Hill, Virginia & Gabriela Alboiu. 2016. *Verb movement and clause structure in old Romanian*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198736509.001.0001>
- Hofmann, Johann Baptist & Anton Szantyr. 1972. *Lateinische Syntax und Stilistik: Mit dem Allgemeinen Teil der lateinischen Grammatik*. Volume 2. Munich: Beck.
- Hooper, Joan B. & Sandra A. Thompson. 1973. On the applicability of root transformations. *Linguistic Inquiry* 4(4). 465–497.
- Jordan, Maria. 2009. Loss of infinitival complementation in Romanian diachronic syntax. Doctoral dissertation, University of Florida.
- Joseph, Brian D. 1983. *The synchrony and diachrony of the Balkan infinitive: A study in areal, general, and historical linguistics*. Cambridge: Cambridge University Press.
- Ledgeway, Adam. 1998. Variation in the Romance infinitive: The case of the southern Calabrian inflected infinitive. *Transactions of the Philological Society* 96(1). 1–61. <https://doi.org/10.1111/1467-968X.00023>
- Ledgeway, Adam. 2005. Moving through the left periphery: The dual complementiser system in the dialects of southern Italy. *Transactions of the Philological Society* 103(3). 339–396. <https://doi.org/10.1111/j.1467-968X.2005.00157.x>
- Ledgeway, Adam. 2012. *From Latin to Romance: Morphosyntactic typology and change*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199584376.001.0001>
- Ledgeway, Adam. 2013. Greek disguised as Romance? The case of southern Italy. In Mark Janse, Brian D. Joseph, Angela Ralli & Metin Bagriacik (eds.), *MGDLT5. Proceedings of the 5th International Conference on Modern Greek Dialects and Linguistic Theory*, 184–227. Patras: Laboratory of Greek dialects, University of Patras.
- Ledgeway, Adam. 2015. Reconstructing complementiser-drop in the dialects of the Salento: A syntactic or phonological phenomenon?. In Theresa Biberauer & George Walkden (eds.), *Syntax over time: Lexical, morphological, and information-structural interactions*, 146–162. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199687923.003.0009>
- Ledgeway, Adam. 2016a. Clausal complementation. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford guide to the Romance languages*, 1013–1028. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199677108.003.0063>
- Ledgeway, Adam. 2016b. The dialects of southern Italy. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford guide to the Romance languages*, 246–269. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199677108.003.0016>
- Leone, Alfonso. 1995. *Profilo di sintassi siciliana. Materiali e ricerche dell'atlante linguistico della Sicilia* 3. Palermo: Centro di Studi filologici e linguistici siciliani.
- Loporcaro, Michele. 1997. Puglia and Salento. In Martin Maiden & Mair Parry (eds.), *The dialects of Italy*, 338–348. London: Routledge.
- Mancarella, Giovan Battista (ed.). 1998. *Salento: monografia regionale della 'Carta dei dialetti italiani'*. Lecce: Del Grifo.
- Meyer-Lübke, Wilhelm. 1899. *Grammatik Der Romanischen Sprachen*, Volume 3. Syntax. Leipzig: Reisland.
- Munaro, Nicola. 2016. Il raddoppiamento del complementatore nelle varietà italo-romanze in prospettiva diacronica. In Jacopo Garzonio & Silvia Rossi (eds.), *Quaderni Di Lavoro ASIt* 19. 89–104.
- Nedelcu, Isabela Mihaela, Adina Dragomirescu, Dana Niculescu & Gabriela Pană Dindelegan. 2016. Non-finite verb forms and non-finite constructions. In Gabriela Pană Dindelegan (ed.), *The syntax of old Romanian*, 232–287. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198712350.003.0003>

- Nicolae, Alexandru. 2015. *Ordinea constituenților în limba română: o perspectivă diacronică. Structura propoziției și deplasarea verbului*. Bucharest: Editura Universității din București.
- Nicolae, Alexandru. 2016. The syntactic configurations of Romanian modal verbs: Modals and phases. Paper presented at MIT LingLunch, Cambridge, MA, May 12.
- Paoli, Sandra. 2003. *COMP and the left-periphery: Comparative evidence from Romance*. University of Manchester. PhD thesis.
- Rivero, Maria Luisa. 1994. Clause structure and V-movement in the languages of the Balkans. *Natural Language & Linguistic Theory* 12(1). 63–120. <https://doi.org/10.1007/BF00992746>
- Rizzi, Luigi. 1997. The fine structure of the left periphery. In Liliane Haegeman (ed.), *Elements of grammar. Handbook on 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, Ian G. & Anna Roussou. 2002. The Extended Projection Principle as a condition on the tense dependency. In Peter Svenonius (ed.), *Subjects, Expletives, and the EPP*, 125–155. Oxford: Oxford University Press.
- Roberts, Ian G. & Anna Roussou. 2003. *Syntactic change: A minimalist approach to grammaticalization*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486326>
- Rohlf, Gerhard. 1969. *Grammatica storica della lingua italiana e dei suoi dialetti*. Volume 3: Sintassi e formazione delle parole. Turin: Einaudi.
- Sandfeld, Kristian. 1930. *Linguistique balkanique: Problèmes et résultats*. 31. Paris: Klincksieck.
- Sgrilli, Paola. 1984. *Il “Libro di Sidrac” salentino. Edizione, spoglio linguistico e lessico*. Pisa: Pacini.
- Sorrento, Luigi. 1951. *Sintassi romanza: Ricerche e prospettive*. (2nd edn.) Varese-Milan: Istituto Editoriale Cisalpino.
- Squillaci, Maria-Olimpia. 2016. When Greek meets Romance: A morphosyntactic analysis of language contact in Aspromonte. Doctoral dissertation, University of Cambridge.
- Stan, Camelia. 2007. Notă gramaticală: “conjuncția” *ca(...)*șă. *Studii și cercetări lingvistice* 58(2). 451–458.
- Stowell, Tim. 1982. The tense of infinitives. *Linguistic Inquiry* 13(3). 561–570.
- Taylor, Cameron. 2014. Aspects of clause structure in Nicotereze. A descriptive account. Doctoral dissertation, University of Cambridge.
- Todorović, Neda & Susi Wurmbrand. 2020. Finiteness across domains. In Teodora Radeva-Bork & Peter Kosta (eds.), *Current Developments in Slavic Linguistics. Twenty Years After*, 47–66. Bern, Switzerland: Peter Lang.
- Vincent, Nigel. 1988. Latin. In Martin Harris & Nigel Vincent (eds.), *The Romance languages*, 26–78. London: Routledge.
- Willis, David. 2007. Specifier-to-head reanalyses in the complementizer domain: Evidence from Welsh. *Transactions of the Philological Society* 105(3). 432–480. <https://doi.org/10.1111/j.1467-968X.2007.00194.x>
- Zafu, Rodica. 2013. Mood, tense, and aspect. In Gabriela Pană Dindelegan (ed.), *The grammar of Romanian*, 24–64. Oxford: Oxford University Press.
- Zafu, Rodica, Alexandru Nicolae, Dana Niculescu, Gabriela Pană Dindelegan, Irina Nicula Paraschiv, Adina Dragomirescu & Andra Vasilescu. 2016. The verb and its arguments: The root clause. In Gabriela Pană Dindelegan (ed.), *The syntax of old Romanian*, 14–231. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198712350.003.0002>

Differential object marking

What type of licensing?

Monica Alexandrina Irimia

University of Modena and Reggio Emilia

Under many recent formal accounts, differential object marking has been taken to signal nominals that must undergo licensing in the clausal syntax, as they bear an [uC] feature (Ormazabal & Romero 2013a; Alcaraz 2018; Bárány 2018; Kalin 2018, among others). While this implementation can capture (standard) Spanish data, the empirical facts from Romanian and Neapolitan I address in this contribution support a view where the differential marker must rather be associated with an additional licensing operation beyond [uC]. More generally, this split appears to be an important locus of parametrization in Romance differential object marking, also confirming similar findings in Ledgeway et al. (2019) for other Romance languages.

Keywords: differential object marking, DP, licensing, animacy, referentiality

1. Differential object marking and licensing

Several Romance varieties are well-known for exhibiting a split in the morpho-syntactic marking of their objects, under an instantiation of the phenomenon called differential object marking (DOM).¹ The examples in (1) from modern Standard Spanish show that animate objects² are signaled by special marking which is homophonous with the dative³ preposition (1a); inanimates, on the other hand, are ungrammatical with the same marker (1b):

1. See especially Bossong (1991, 1998); Torrego (1998); Aissen (2003), for a few landmark titles, as well as López (2012) for an extensive list of references.

2. Under specified syntactic conditions, as I show later in the chapter.

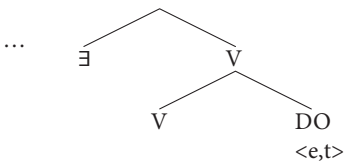
3. I do not address whether DOM must be structurally unified with DAT (see Manzini & Franco 2016 for a positive answer). My data strongly indicate that DOM has accusative syntax (in Romanian DOM and DAT are not homophonous; in Neapolitan DOM shows agreement characteristic of accusatives, and in both these languages, clitic doubling of DOM can only take ACC morphology, etc.).

- (1) a. *He encontrado *(a) la niña*
 have.1SG found DAT=DOM the girl
 “I have found the girl.”
- b. *He encontrado *(a) el libro*
 have.1SG found DAT=DOM the book
 “I have found the book.”

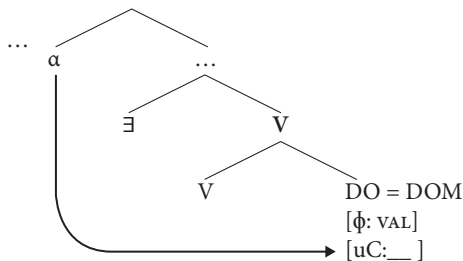
(Spanish, Ormazabal & Romero 2013a: Example 1a, b)

Recent formal research has provided a new perspective on the nature of this type of morphology. Starting mainly from (standard) Spanish data, contributions by Rodríguez-Mondoñedo (2007), Ormazabal & Romero (2013a; b), Alcaraz (2018); Kalin 2018 or Bárány (2018), among others, associate DOM to those nominals that require licensing in the syntax. The latter contain an uninterpretable Case ([uC]) feature and must enter into an adequate checking relationship with a functional head in the clausal spine (3). Inanimates as in (1b), on the other hand, are unlicensed (2); in fact, glossing over differences between these accounts, they are also considered to be predicates (of type $\langle e, t \rangle$), found under existential closure (\exists) and undergoing semantic (pseudo-)incorporation with the main verb (see especially Alcaraz 2018 or Bárány 2018, among others).

- (2) Non-differentially marked objects



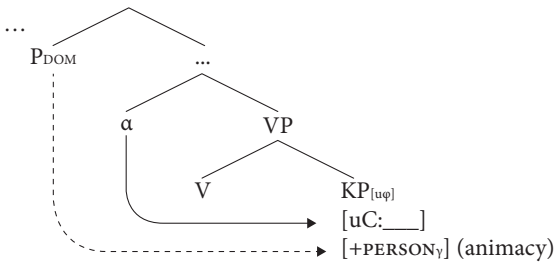
- (3) Differentially marked objects – standard Spanish



This short contribution addresses further refinements on the licensing accounts. The main conclusion I support is that DOM is not only restricted to the split between Case-licensed and unlicensed nominals. An examination of two other robust DOM languages, namely Romanian and Neapolitan, indicates that the prepositional

marker (also referred to as oblique DOM), in fact, tracks an additional licensing operation on objects that might have already checked their [uC] feature. Thus, both Romanian and Neapolitan exhibit objects of the type in (2) and (3); however, some classes of nominals (among which DOM animates, under certain conditions) also contain a [+PERSON_γ] feature linked to animacy (γ from gender to distinguish it from other [PERSON] features, see § 3.1 and § 4), which enters into an additional licensing operation, as shown in the simplified representation in (4). This difference forms the basis of a micro-variation point in Romance DOM, as also independently noticed in the recent discussion by Ledgeway et al. (2019).

(4) DOM as additional licensing – Romanian and Neapolitan



The structure of the chapter is as follows. In § 2 I discuss data from modern Spanish varieties which have prompted an analysis of differential objects as licensed nominals, as opposed to the unmarked types which are assumed to be subject to incorporation. In § 3, I introduce the (relevant) facts from Romanian and outline their differences from Spanish; I motivate the assumption that Romanian DOM should be equated with an additional licensing mechanism (see also Cornilescu & Tigău 2017). In § 4 I show that the same supplementary licensing account is also necessary for Neapolitan DOM. Section 5 contains the conclusions.

2. Differential object marking in Spanish

Grammars of Spanish, especially descriptive ones, often remark that animate objects, if definite, are not possible without oblique DOM. Thus, examples such as (1a) are generally taken to be ungrammatical without the dative preposition. However, when tested more carefully, native speakers do accept examples similar to (5), where a definite animate is used without the oblique differential marker; crucially, the only possible interpretation is that of a non-specific and/or non-referential definite. This seems to match the facts grasped from indefinites; as seen in (6), if the differential marker is absent, only a non-specific reading is possible, while indefinites with the

prepositional accusative (6) can easily get specific readings.⁴ The same facts hold with other non-specific definites, for example Quine-type definites illustrated in (7); for many speakers these are only possible without differential marking, although the nominal is both animate and definite.

- (5) *He encontrado la niña que buscas*
 have.1.SG found the girl that search.2SG
 “I have found the type of girl you are looking for.”
- (6) a. *He encontrado una niña*
 have.1SG found a girl
 “I have found some girl or other.”
 b. *He encontrado a una niña*
 have.1SG found DAT=DOM a girl
 “I have found a (specific) girl.”
- (7) *Juan busca la mujer perfecta*
 Juan search.3.SG the woman perfect
 “Juan is looking for the perfect woman.” (Quine definite)

Given that non-specific indefinites, as well as (Quine-type) non-referential definites normally require an analysis as predicates of type $\langle e,t \rangle^5$ (as opposed to being nominals with argumental nature), these examples have been taken to indicate that Spanish definiteness morphology does not introduce referentiality/specificity by default and moreover can be of type $\langle e,t \rangle^6$ (see especially Rodríguez-Mondoñedo 2007; López 2012 or the extreme view in Alcaraz 2018). One of the results is that in the syntax, Spanish definite non-specific animates do not need licensing in the same way as the DOM-ed specific animates, as they do not contain a [uC] feature (see also Bárány 2018).

These assumptions are also strengthened by an examination of the syntactic configurations where differential marking is, in fact, not possible. Ormazabal & Romero (2013b) discuss contexts containing an indirect object which is doubled

4. Although the differential marker is in no way restricted to specificity. The same conclusion results from López (2012), who provides examples containing DOM (in)definites as well as the subjunctive mood, an unambiguous signal of non-specificity. I illustrate some of these throughout the paper, see especially (10) and (11).

5. See also recent discussion by Espinal & Cyrino (2017) about other types of non-specific definites in Romance, which also need an account in terms of predicates.

6. As surprising as it might be, the conclusion that definiteness morphology (or more generally, D^0) always introduces a $\langle e,t \rangle$ category (creating predicates) has also been claimed for other languages with rich DPs. See, for example, the discussion in Davis (2018), as applied to Salish, specifically Lillooet.

by a dative clitic as in (8). Here DOM results in ungrammaticality and must be removed, even if the direct object is animate and definite.

- (8) *Le enviaron (*a) todos los enfermos a la doctora van Tan.*
 CL.DAT.SG sent.3PL DOM all the sick people DAT the doctor van Tan
 Intended: “They sent all the sick people to doctor van Tan.”
 (Spanish, Ormazabal & Romero 2013b Example (2b))

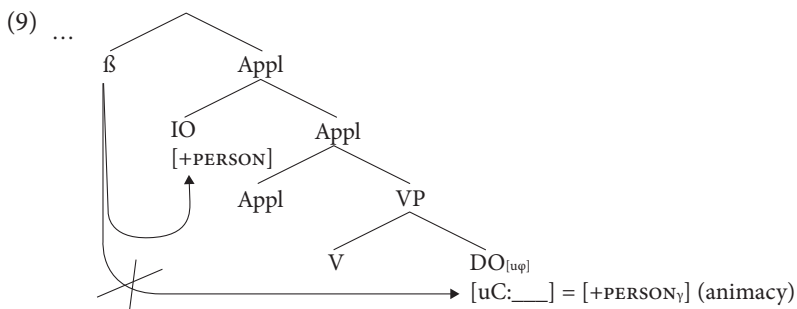
However, as Ormazabal & Romero (2013b) observe, the removal of the differential marker has consequences on the nominal interpretation:

the availability of (8) [O&R’s 2b, my note] is extremely restricted. Sentences like (8) [O&R’s 2b, my note] are only grammatical with nouns such as sick people, soldiers, slaves, kids, etc.; nouns whose referents are regularly treated as entities lacking free will. The range of animate nouns that can appear without DOM in this context is, more or less, the same one that allows **incorporation** in polysynthetic languages... (Ormazabal & Romero 2013b: 157)

2.1 Differential object marking as licensing

Examples like (8) provide other non-trivial hints into the nature of DOM. Given that they result in ungrammaticality only when the indirect object is clitic doubled, one possible analysis we entertain here is that they enter into a p(erson) c(ase) c(onstraint)-type competition in this environment. We can make the plausible assumption that dative clitic doubled nominals signal the presence of [+PERSON], which requires licensing in the syntax, by entering into the checking relationship with a relevant licenser endowed with a [+PERSON] probe (Anagnostopoulou 2003; Béjar & Rezac 2003, among others). The ungrammaticality of DOM could then follow from its also being specified with a [+PERSON]_γ feature (interpreted as animacy) which needs licensing. As a given domain typically contains only one licenser with a [+PERSON] probe (Anagnostopoulou 2003, among others), the ungrammaticality of both DOM and a clitic doubled indirect object can be straightforwardly derived: they both compete for the same licenser (9).⁷

7. This is an oversimplification; the analysis as it stands does not explain why ungrammaticality does *not* arise when DOM co-occurs with a dative clitic, but without the overt indirect object DP. In the next sections I discuss data under which DOM, in fact, signals an additional licensing operation; thus, initial licensing operations also seem to count, although the facts are not transparent in Spanish. In any case, the presence of a [+PERSON] feature connected to DOM is necessary.




Additional confirmation that DOM indicates nominals that must undergo licensing in the syntax comes from clause-union contexts. As López (2012) has discussed, in these contexts, the prepositional marker is *obligatory* on animates, irrespective of whether the animate is interpreted as specific or not. We illustrate two examples with small clause (sc) nominals in (10) and (11); note that DOM is obligatory on these animate indefinites, despite their receiving a non-specific interpretation, cf. (6). In (10), the subjunctive mood introduces non-specificity, indicating that *Juan* does not have a specific man in mind that he would consider dishonest. Similarly, in (11), the intensional adjective *necessary* normally triggers non-specific interpretations on its argument. Thus, the most prominent reading is that the professor considers some student or other necessary for the project. These contexts also clearly demonstrate that oblique DOM is *not* a specificity mechanism.

- (10) *Juan non considera honrado* *(a) *un hombre que acepte sobornos*
 Juan not considers honest DOM a man that accept.SBJV bribes
 “Juan does not consider honest a (any) man that would accept bribes.”
 (López 2012: 25 Example 68)
- (11) *El profesor considera* *(a) *un estudiante necesario para el proyecto.*
 the professor considers DOM a student necessary for the project
 “The professor considers a student necessary for the project.” (Spanish)

sc configurations are important as they signal the need for licensing on their nominal. Given their syntactic structure as complements to V, scs do not permit the option of the nominal to undergo (pseudo-)incorporation (see also López 2012); normally, the latter operation can only take place in complement position.

$$(12) \quad [{}_{VP} EA \nu [{}_{AP} \text{NomP } \alpha [{}_{VP} V [{}_{SC} [t(\text{NomP}) AP]]]]$$



No incorporation

Most models of nominal syntax permit just two options: either (pseudo-) incorporation (which can manipulate predicative categories of type $\langle e, t \rangle$ and caseless nominals), or Case licensing via valuation by a relevant functional head on the clausal spine (see also Levin 2015, among others). The sc nominal is only left with the option of licensing in the syntax, and must contain relevant pieces of structure that can host structural [uC] (given that the structural conditions for incorporation cannot be met, as shown above). Therefore, (in)definite animates are predicted not to be possible without oblique differential marking, irrespective of specificity. As we have seen, this is borne out for Spanish; thus, the conclusion emerges that differentially marked objects in Spanish are those categories that carry an uninterpretable Case feature, possibly connected to the presence of [+PERSON_y]. Objects that do not take the differential marker are unlicensed, even if definite.⁸ The table below summarizes the results of this section, also cross-referencing the relevant examples. (√ indicates presence, while – indicates absence of a certain property).

Table 1. Incorporation and licensing in Spanish

	No licensing	Obligatory licensing in the syntax (differential marking)
Specific animates	–	√ (10), (11)
Non-specific animates	√ (6), (8)	–
Definiteness morphology	√ (5), (7)	–
[+PERSON]	–	√ (10), (11), (8), (9)
Animates in scs	–	√ (10), (11)

8. Note however that a binary system of this type will leave some questions unanswered, especially with respect to inanimates. Inanimate definite objects, which do not normally take DOM (leaving aside here dedicated contexts in which they do, as they are orthogonal to the discussion) are not necessarily interpreted as non-specific, across the board. They are also subject to passivization, and (for many speakers) possible in sc contexts even without DOM, indicating that they are active in the syntax. This raises the question about how differences from languages like Romanian are to be derived.

3. Differential object marking in Romanian: Additional licensing

The dichotomic picture (licensing vs non-licensing) illustrated for Spanish is however harder to maintain for other Romance varieties. We discuss here some facts from Romanian, a language with robust oblique DOM, signaled by the preposition *pe*.⁹ The diagnostics we examine below indicate that Romanian DOM rather signals an additional licensing operation, independent of [uCase].

3.1 Types of D⁰ in Romanian

As opposed to Spanish, Romanian animate definites are possible without oblique differential marking (see also Cornilescu 2000; Tigău 2011, among others) and are not restricted just to non-referential/non-specific readings. The non-DOM definite (on an animate noun) in (13) receives a referential, specific interpretation, like the DOM definite in (14):

- (13) *Am văzut fata frumoasă*¹⁰
 have.1 seen girl.the beautiful
 “I/we have seen the beautiful girl.”
- (14) *Am văzut-o¹¹ pe fata frumoasă*
 have.1 seen-CL.F.3SG.ACC DOM girl.the beautiful
 “I/we have seen the beautiful girl.”

For many speakers, these two sentences appear to be interchangeable (but see Cornilescu 2000). They also pose a non-trivial challenge; one cannot simply assume that differential marking is *optional* in Romanian. Like in Spanish, there are systematic contexts where the omission of DOM results in severe ungrammaticality. For lack of space I only provide here an example with the animate negative quantifier:

- (15) *Nu ai văzut *(pe) nimeni*
 not have.2SG seen DOM nobody
 “You (sg.) have not seen anybody.”

Crucially for my purposes, there is also solid indication that non-DOM definites as in (13) are subject to licensing. This is another difference from Spanish. A strong

9. Which also has independent locative uses (“on”).

10. I use adjectival modification on the noun to avoid the problem of definiteness deletion in the complement position of a preposition in Romanian.

11. As opposed to standard Spanish, Romanian DOM nominals can be clitic-doubled. See also § 4.

argument comes from clause-union contexts. I have concluded above that Spanish scs require DOM if the object is animate. Romanian DOM is not obligatory with animates in scs. I present here the Romanian example that corresponds to the Spanish one in (11):

- (16) *Profesorul consideră (pe) un student necesar pentru proiect.*
 professor.the considers DOM a student necessary for project
 “The professor considers a student necessary for the project.”

There is indication that Romanian scs are contexts where nominals indeed require licensing. For instance, nominals are not allowed to occur bare under scs in Romanian, as seen in (17):

- (17) **Profesorul consideră student/studenti inteligent/ inteligenti*
 profesor.the considers student/students intelligent.SG intelligent.PL

Note that this restriction cannot be attributed to the general ban on bare (singular) nouns in object position. The confirmation comes from the behavior of mass nouns; these classes can appear bare when used as objects, irrespective of whether the predicate is intensional or not.

- (18) *Caută/cumpără/consideră miere (pentru recepție)*
 searches/buys/considers honey for reception
 “S/he is looking for/buys/considers honey (for the reception)”

When used in scs, mass nouns are ungrammatical in their bare form – definiteness morphology is obligatory:

- (19) *Consideră *miere/√ miera sănătoasă*
 considers honey/honey.the healthy
 “S/he considers (the) honey healthy.”

I connect this restriction to the need of nominals in scs to undergo licensing, given that they cannot undergo (pseudo-) incorporation, as discussed above.¹²

12. Note that associating the presence of overt determiners to the (putative) subject status of the nominal inside the sc (under the observation that subjects are not normally determiner-less across Romance) does not seem to be on the right track. On the one hand, under many accounts, *consider*-type adjectival structures do not have clausal status, but rather form complex predicates to which the shared argument is merged compositionally (see Williams 1983, among others). On the other hand, even if the sc structure is maintained, it appears not to be a domain of quantification, and thus determiner layers could not merge inside it (see Moulton 2013 for recent discussion). There are many other empirical and theory-internal arguments against the sc subject hypothesis, e.g. subjects can appear bare, at least in certain configurations.

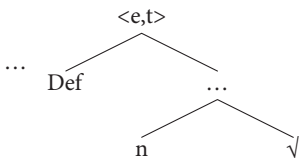
More precisely, these facts can be straightforwardly derived under the assumption that the functional head D^0 hosts a [uC] feature in Romanian (see also Giusti 1993), which requires licensing and cannot undergo incorporation. Thus, sc nominals always need to show overt morphology signaling the presence of the DP layer. This indicates a locus of microvariation. In Spanish, the default D^0 introduces/outputs categories that can be left unlicensed. In Romanian, on the other hand, D^0 normally requires licensing, irrespective of the presence of additional quantificational or Case bearing material. The differences are summarized in Table 2:

Table 2. Types of D^0 and licensing

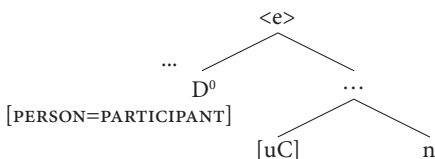
D^0	No licensing	Licensing ([uC])
Spanish	√ (5), (7)	-
Romanian	-	√ (13), (16), (17)

A tension between two (opposite) views on the (semantic) nature of definiteness morphology can also be reconciled: (a) as a predicative category, as outlined above and in fn. 6; (b) definiteness in D^0 as argumental (of type $\langle e \rangle$, as discussed in contributions by Longobardi (2008); Longobardi & Guardiano (2009), among others, and requiring licensing). More specifically, in the latter view, D^0 contains a [PERSON] feature which contributes argumenthood, or the presence of a category with a [PARTICIPANT] value. Data from Romanian (and from Neapolitan, as we will see in the next section) show that the two possibilities are instantiated in the same language. We schematize this in (20):

- (20) a. Non-referential (predicative) definiteness



- b. Referential D^0 (following Giusti 1993; Longobardi 2008, among others)



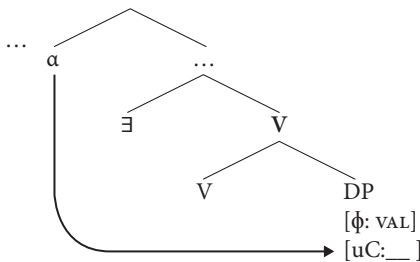
Like in Spanish, some Romanian definites only get a non-specific, non-referential reading, even if the DP is animate. In this case, differential marking is *not* possible on the definite object. Among these are the Quine-definites exemplified for Spanish in (7) above. In (21) a non-specific, non-referential reading can only be obtained if the definite does not have differential marking.¹³ The addition of differential marking in (22) entails referentiality, but not necessarily specificity/familiarity.¹⁴

- (21) *Ion (încă mai) caută (*pe) femeia perfectă*
 Ion still more searches DOM woman.the perfect
 “Ion is (still) looking for the perfect woman.”
 (under a Quine non-specific, non-referential reading)

- (22) *Ion (încă) o (mai) caută pe femeia perfectă*
 Ion still CL.3F.SG.ACC more searches DOM woman.the perfect
 “Ion is (still) looking for the perfect woman.” (under a referential reading)

However, based on the (sc) evidence discussed in this subsection, it is safe to conclude that Romanian non-differentially marked definites can also undergo licensing as they are associated with [uC]¹⁵ when in D⁰, irrespective of the presence of other pieces of structure that might require (additional) licensing. The schematic representation is as in (23). But this leaves unanswered the question of how DOM is to be analyzed.

- (23) DP (definiteness in D⁰, etc.) licensing in Romanian



13. And in these structures the definiteness morphology is probably not merged in D⁰ (but in a lower position), but ends up there after raising with N.

14. It might be the case that Ion or the audience might not know precisely who the right individual is who qualifies as the perfect woman. The minimum condition here is that the perfect woman must at least have been introduced in the previous discourse or is part of a set that has been made known and actualized.

15. See also Cornilescu & Tigău (2017), who motivate the same conclusion for Romanian.

3.2 Differentially marked objects in Romanian

DOM shares important commonalities in Romanian and Spanish. First, in extensional contexts as in (1), it can only affect animates. Second, there is evidence that the DOM preposition signals the presence of a [+PERSON] feature. We have seen examples from Spanish PCC-type effects in (8) and (9); in Romanian, the correspondent of (8) is also degraded, especially if both the clitic-doubled indirect object and DOM are plural:¹⁶

- (24) **Le-au trimis pe toți bolnavii doctorilor.*
 CL.3PL.DAT-have.3.PL sent DOM all sick-people.the doctor.PL.DAT
 Intended: “They have sent all the sick people to the doctors.”

Another well-known restriction in Romanian affects the co-occurrence of DOM with possessor raising.¹⁷ The sentence in (25) is simply ungrammatical for many speakers; the differential marker must be removed to ensure grammaticality as in (26):

- (25) **Ion și (l)-a văzut pe un prieten.*
 Ion CL.3SG.DAT CL.3M.SG.ACC-has seen DOM a friend
lit. “Ion to himself saw a friend.”
 Intended: “Ion saw a friend of his.”
- (26) *Ion și-a văzut un prieten.*
 Ion CL.3SG.DAT-has seen a friend
 “Ion saw a friend of his.”

As mentioned above, I follow a large body of work which connects animacy (as spelled-out by the oblique differential marker) to the presence of [+PERSON_γ] feature (Cornilescu 2000; Rodríguez-Mondoñedo 2007; Richards 2008, among others), merged on a gender-introducing projection inside the nominal, hence the diacritic γ . The schematic picture obtained is in the table below, adapted from Harley & Ritter (2002) or Nevins (2007):¹⁸

16. The example is equally degraded if clitic doubling on DOM is added. See also Cornilescu & Tigău (2017) for other examples. However, note that the sentence is well-formed if both the clitic-doubled IO and DOM are in the singular, or if they do not match in number. Such refined restrictions are not surprising for DOM, and are also found outside Romance; however, an exhaustive discussion would go beyond the limits of this paper.

17. This is not unambiguous evidence that DOM is structurally DAT (fn. 3), as there are also contexts where ACC clitic-doubling is not possible with DOM. See the previous footnote.

18. I use a binary specification for [PERSON] just for convenience. Other implementations are, of course, possible.

Table 3. Person and animacy (building on Harley & Ritter 2002)

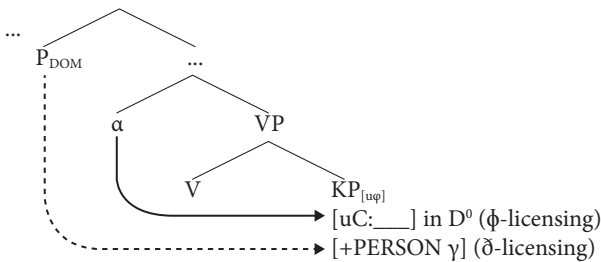
PERSON/ANIMACY	FEATURES
1st person	[+PERSON] (= [+PARTICIPANT]) speaker
2nd person	[+PERSON] (= [+PARTICIPANT]) addressee
3rd person	[+PERSON] _v (= [+PARTICIPANT])
[+human, +animate]	

Now, the problem that arises is the following: I have mentioned that, for authors like Longobardi (2008) or Longobardi & Guardiano (2009), semantically strong/referential D^0 is also associated with [+PERSON], interpreted as [+PARTICIPANT], irrespective of animacy. However, as we have seen above, definiteness does not enter into PCC effects. This indicates that it does not require licensing from the same licenser as the dative clitic double. On the other hand, assuming that both DOM as well as the dative-clitic-doubled indirect object compete for [PERSON] licensing is a good start to explaining the ‘PCC-effects’.

The only way to reconcile these two observations is that there can be various types of [+PERSON] features in the make-up of a nominal (see also Bárány 2017). How are they to be set apart formally?

I propose that part of the answer resides in the modality of licensing these various types of [+PERSON] features are subject to. Following recent insights in Miyagawa (2010, 2017), I make a crucial distinction between ‘grammatical’ (ϕ) licensing, and ‘discourse’ (δ) licensing. The strongest type of nominal licensing is δ licensing. Nominals that bear the relevant discourse-salient features must be licensed by functional projections that are linked to the Anchoring Projection, the highest functional domain in the upmost periphery of the clause. There are also nominals that require licensing in narrow syntax, but via a strictly grammatical procedure, and whose licensers are not necessarily linked to the Anchoring Domain. Nominals with ([+PERSON]) can undergo weakening in that they can lose their δ -feature and might require just (possibly fallible) ϕ -licensing (associated with [uC] $_{\phi}$, tracking gender and number as overtly seen for Neapolitan in § 4), or might lose their need for licensing altogether, being always predicates and subject to incorporation. I have reviewed accounts proposing that definiteness morphology in Spanish does not come with an obligatory licensing requirement. Romanian, on the other hand, preserves a type of [+PERSON] feature, which only requires [uC] $_{\phi}$ -licensing. [+PERSON]_v, on the other hand, appears to always be connected to a δ -licensing procedure in the two languages, must always enter into checking relations in sentential syntax, and cannot be subject to incorporation. Romanian, thus, provides evidence that DOM undergoes an additional licensing operation, namely the one connected to δ -licensing (27). This also accounts for the observation that DOM requires an object of DP-type, and not a bare nominal or a predicative category.

(27) Romanian DOM-licensing



To repeat, we have examined evidence that Romanian nominals exhibit (at least) two types of licensing in the syntax. As opposed to Spanish, Romanian DOM must be understood as an additional licensing operation which tracks the presence of $[+PERSON_\gamma]$, associated to δ -licensing. Table 4 summarizes the differences we have investigated for Spanish and Romanian.

Table 4. Licensing in Spanish and Romanian

	Unlicensed	Obligatory differential marking
Specific animates	Spanish – Romanian –	Spanish \checkmark (10) Romanian – (13)
Non-specific animates	Spanish \checkmark Romanian –	Spanish – (7) Romanian – (21)
Definiteness in D^0	Spanish \checkmark Romanian –	Spanish – (5) Romanian – (13)
$[+PERSON_\gamma]$ (discourse-salient animacy)	Spanish – Romanian –	Spanish \checkmark (1a) Romanian \checkmark (14)
Animates in scs	Spanish – Romanian –	Spanish \checkmark (10), (11) Romanian – (16)

4. Differential object marking in Neapolitan: More evidence for additional licensing

Neapolitan is another Romance variety that shows an object split broadly regulated by animacy. The morphological means are a preposition that is homophonous with the dative (the *a* marker). As we can see in (28), the special animate *aragosta* (“lobster”) can be differentially marked. Thus, Neapolitan patterns with Romanian, and is unlike Spanish, in the sense that animate referential/specific definites can be used without oblique DOM in a wide variety of contexts (see also Ledgeway 2000, or Ledgeway et al. 2019, among others):

- (28) (L')*addʒə* **kwottə*/√*kəttə* (*a*) *l'aragosta*
 CLT.3F.SG.ACC-have.1 cooked.M.SG/F.SG DOM the.F.SG -lobster
 “I have cooked the lobster.” (Adam Ledgeway, Roberto Petrosino, p.c.)

Neapolitan, however, raises another very interesting problem. If the example above is carefully analyzed, one also notices the presence of overt object agreement. More precisely, the auxiliary shows agreement with the subject (a first person agent), while past participle agreement (PPA)¹⁹ is always with the object. Crucially, PPA is independent of DOM; in (29) we illustrate a context, originally provided by Loporcaro (1998: 68–69),²⁰ where PPA tracks an inanimate definite object, which cannot take DOM:

- (29)
- | | | | | |
|--------------|--------------------------------|--|------------------------|--------------|
| <i>addʒə</i> | <i>*kwottə</i> /√ <i>kəttə</i> | | <i>a</i> ²¹ | <i>pastə</i> |
| have.1 | cooked.M.SG/F.SG | | the.F.SG | pasta |

A further remark is necessary; the pioneering discussion in Loporcaro (1998, 2010) has clearly demonstrated that PPA does not have to always be associated with overt object movement, contrary to classic Kaynean analyses. We have clear proof for this assumption in these Neapolitan examples, where objects trigger agreement, although they are not found in a dislocated position. This picture is also common across other southern Italian varieties, and also outside Italy (see Loporcaro 1998, 2010 for further exemplification).

19. PPA is only seen with those predicates that exhibit a root-internal change known as ‘metaphony’ (see Loporcaro 1998, 2010, among others).

20. And confirmed by the native speakers consulted here. See also Ledgeway (2000: 306).

21. The differential marker and the feminine definite article are homophonous in Neapolitan nominal roots that start in a consonant. We know however that the *a*-marker in (29) is not DOM as it shows up on an inanimate nominal.

What matters for our present investigation is that the derivation of PPA requires a licensing account in Neapolitan; thus, we are left with the same puzzle as for Romanian – what type of mechanism is involved in oblique differential marking? One possibility that comes to mind is to associate object agreement with a ‘lexical’ or ‘inherent’ process, as has been claimed by Rodríguez-Mondoñedo (2007) for dialectal Spanish data.²² The biggest problem such assumption would encounter in Neapolitan comes from accusative clitic-doubling. I have noted that definite animates are possible without differential marking, examples like (30) being well-formed. We notice here, however, that accusative clitic doubling is also *possible*, despite overt PPA. Neapolitan thus differs from standard Spanish, where accusative clitic doubling is ungrammatical with non-pronominal DPs, regardless of differential marking. But if we wanted to propose a lexical/inherent mechanism for PPA, it is not clear how clitic-doubling would be derived. Under most accounts, accusative clitic doubling is a structural licensing mechanism.

- (30) *L'addzə* **kwottə*/√*kəttə* *l'aragostə*
 CLT.3F.SG.ACC-have.1 cooked.M.SG/F.SG the.F.SG -lobster
 “I have cooked the lobster.”

Thus, Neapolitan PPA must be structural; it also guides us towards elucidating the nature of oblique DOM. As we have mentioned, PPA tracks (referential) direct objects, irrespective of animacy. DOM, on the other hand, only affects a subclass of PPA objects, generally the animates. Corroborating this with the observation that Neapolitan DOM also triggers PCC like effects similar to in Spanish or Romanian,

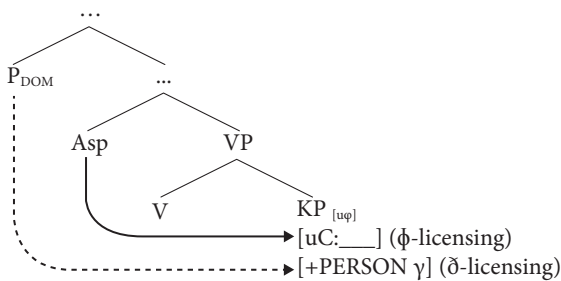
22. One piece of evidence for the lexical/inherent marking hypothesis is that, in the variety Rodríguez-Mondoñedo (2007) discusses, object agreement extends to existential clauses (ECs). As we see in (ib), the clitic carries accusative morphology (indicating its direct object nature), but still triggers object agreement on the existential. As ECs are normally assumed to be domains where structural Case on the object does not obtain (given that the nominal rather behaves like a predicate, blocking referential definiteness or specificity), this motivates an inherent/lexical approach to object agreement. However, there are several problems with extending this assumption to other Romance varieties. As is well known, at least since Zamparelli (2000), ECs can have complex structure in Romance – for example, overt definites are possible in Italian and other varieties, etc. Thus, the inherent hypothesis cannot be extended to all Romance without carefully investigating the type of ECs or D⁰ present in any given variety.

- (i) a. *Hubieron dos estudiantes en la fiesta.*
 be.PST.3PL two students at DEF.F.SG party
 “There were two students at the party.”
 b. *Los hubieron.*
 CL.ACC.3PL be.PST.3PL
 “They were.” (Rodríguez-Mondoñedo 2007, Example 2)

and thus signals a [+PERSON_γ] feature which must be δ -licensed, the only option left for us is that DOM, once again, signals an additional licensing operation. The same conclusion was reached recently by Ledgeway et al. (2019). Additionally, using canonical c-command tests (see López 2012), we also know that binding into the E(xternal) A(rgument) is not possible from objects that show PPA, or that are involved in DOM, if they are not clitic-doubled. This provides evidence that both PPA and oblique DOM objects are generated and interpreted lower than the EA.

We follow D’Alessandro & Roberts (2010) in assuming that PPA is connected to a low Asp head merged immediately above V. This enters into a ϕ -relationship with the object, checking the latter’s [uC]. The [+PERSON_γ] feature will need an additional licenser,²³ which, I propose (following López 2012), is an intermediate functional projection between VP and ν .

(31)



Importantly, at least some types of clitic doubling on the (DOM-ed) object allow binding into EA to go through. Extensive homophony between the ACC and DAT forms of the clitic might make the testing of the relevant contexts non-transparent in Neapolitan. However, in Romanian, where ACC morphology is more clearly individuated for clitics, ACC clitic doubling of DOM allows binding into the EA, as opposed to the non-clitic doubled DOM contexts. The relevant contrast is below, between (32) and (33). That the ACC clitic makes available an antecedent for binding (in a position above the EA) has been observed for other languages (see also Anagnostopoulou 2006):

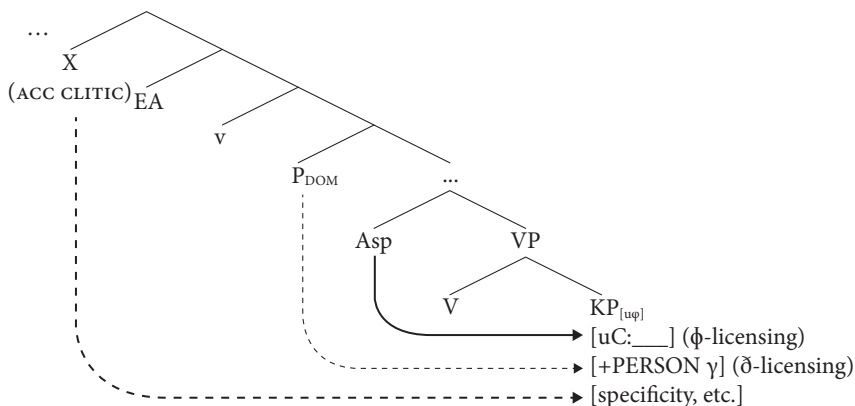
- (32) *Părinții lor_{i,j} i_i-au lăudat pe mulți_i copii*
 parents.the their CL.ACC.3M.PL-have.PL praised DOM many children
 “Their (own) parents have praised many children.”

23. Also note that PPA only tracks number and gender. This indicates that Asp does not contain a [+PERSON] probe.

- (33) *Părinții lor^{*i/j} au lăudat pe mulți_i copii*
 parents.the their have.PL praised DOM many children
 “Their parents have praised many children.”

We can conclude from these contrasts that there is an additional licensing locus for objects above vP^{24} (34), which is generally associated with specificity (in the languages discussed here). Obviously, much more needs to be said about clitic doubling and its interactions with both DOM, as well as non-DOM DPs.²⁵ My goal here was simply to show that DOM is by *no means* the only structural licensing strategy available for objects across Romance.²⁶ In Neapolitan, PPA, DOM and ACC clitic doubling are independent licensing mechanisms, despite non-trivial interactions, and similar semantico-pragmatic outputs. The same observation holds for Romanian DOM and ACC clitic doubling, as has also been argued for extensively in Cornilescu (2000).

(34) Object licensing in Romance



24. In both Romanian and Neapolitan, clitic doubled objects (DOM-ed or not, for Neapolitan) cannot easily escape specificity readings. We have seen throughout this chapter that DOM itself is not intrinsically connected to specificity.

25. And why, in some contexts clitic doubling is obligatory, while in others it results in ungrammaticality. Also, the question of whether the feature associated with accusative clitic doubling is the same as the one related to [uC] on D^0 or yet a different ([PERSON]) feature is by no means trivial. However, the space limitations do not allow us to address all these aspects in detail.

26. Interestingly, Wiltschko (2014) also discusses evidence for three structural licensing loci for objects, with respect to Algonquian.

To conclude this final section, Neapolitan, like Romanian, provides evidence that DOM rather signals an additional licensing operation on certain classes of objects. The main difference between these two languages and standard Spanish is that their D^0 contains a [uC] feature that requires licensing. This [uC] feature is independent of DOM. This microvariation point has non-trivial consequences on the licensing modalities of these languages, and on the conditions of realization of DOM itself (Table 5). The results obtained in this paper are summarized in (Table 6).

Table 5. Types of D^0 in Romance

Definiteness in D^0	Unlicensed	Licensing ([uC])
Spanish	√ (6)	-
Romanian	-	√ (16), (17)
Neapolitan	-	√ (28), (30)

Table 6. Incorporation and licensing in Spanish, Romanian and Neapolitan

	Unlicensed	Obligatory differential marking
Specific animates	Spanish -	Spanish √ (1a), (6)
	Romanian -	Romanian - (13)
	Neapolitan -	Neapolitan - (28)
Non-specific animates	Spanish √	Spanish - (5), (7)
	Romanian -	Romanian - (21)
	Neapolitan -	Neapolitan -
Definiteness in D^0	Spanish √	Spanish - (5)
	Romanian -	Romanian - (13)
	Neapolitan -	Neapolitan - (30)
[+PERSON _y]	Spanish -	Spanish √ (1a)
	Romanian -	Romanian √ (14)
	Neapolitan	Neapolitan √ (28)
Animates in scs (irrespective of specificity)	Spanish -	Spanish √ (10), (11)
	Romanian -	Romanian - (16)
	Neapolitan -	Neapolitan -

5. Conclusions

The (limited set of) Neapolitan and Romanian data analyzed here indicate more transparently that objects can be structurally licensed in a variety of positions inside ν P and above ν P (via clitic-doubling). DOM is just one of these strategies (see also Suñer 1988). An important question is whether Standard Spanish DOM could also be equated with additional licensing, despite superficial evidence to the contrary (see also fn. 8). A detailed examination of inanimates could provide support in this direction: under *scs* (which block object incorporation), DP morphology is obligatory similarly to in Romanian and Neapolitan. This might suggest that Spanish non-DOM D^0 could, in fact, undergo licensing under certain conditions, supporting an analysis along the lines proposed by López (2012). This leaves the additional licensing for DOM as a possible option, with the added constraint of a licensing competition between DOM and D^0 in contexts involving referential animates. However, there also seem to be speakers who prefer to extend the differential morphology to all types of nominals in *scs*, irrespective of animacy. A much more detailed investigation of inanimates is nevertheless necessary in order to fully spell-out these mechanisms. Another aspect that requires further extensive investigation is the connection between the additional licensing account proposed here for Romanian and Neapolitan and analyses that link DOM to information-structure requirements (e.g., topicality).

Acknowledgements

I would like to thank Alexandra Cornilescu, Anna Pineda, Giuseppe Longobardi, Adam Ledgeway, Roberto Petrosino, as well as the audience at Going Romance 31 (2017, Bucharest) for discussion, data, and valuable suggestions. I am also grateful to Adina Dragomirescu and Alexandru Nicolae for their extremely useful feedback and hard work throughout the editing process. All errors are my own.

Abbreviations

ACC	accusative	PCC	Person Case Constraint
CL	clitic	PL	plural
DAT	dative	PPA	past participle agreement
DOM	differential object marking	PST	past
EA	external argument	SC	small clause
EC	existential clause	SBJV	subjunctive
F	feminine	SG	singular
M	masculine	1/2/3	person
NOMP	nominal phrase		

References

- Aissen, Judith L. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21(3). 435–483. <https://doi.org/10.1023/A:1024109008573>
- Alcaraz, Alejo. 2018. Deep and surface clitics in northern Castilian Spanish. Paper presented at GLOW 2018, Budapest, April 10–14.
- Anagnostopoulou, Elena. 2003. *The syntax of ditransitives: Evidence from clitics*. Berlin: Mouton de Gruyter.
- Anagnostopoulou, Elena. 2006. Clitic Doubling. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax* (1st edn.), Volume 1, 519–581. Malden, MA: Blackwell. <https://doi.org/10.1002/9780470996591.ch14>
- Bárány, András. 2017. *Person, case and agreement: The morphosyntax of inverse agreement and global case splits*. Oxford: Oxford University Press.
- Bárány, András. 2018. DOM and dative case. *Glossa: A Journal of General Linguistics* 3(1). 1–40. <https://doi.org/10.5334/gjgl.639>
- Béjar, Susana & Milan Rezac. 2003. Person licensing and the derivation of the PCC effects. In Ana Theresa Pérez-Leroux & Yves Roberge (eds.), *Romance linguistics: Theory and acquisition*, 49–62. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.244.07bej>
- Bossong, Georg. 1991. Differential object marking in Romance and beyond. In Dieter Wanner & Douglas A. Kibbee (eds.), *New analyses in Romance linguistics*, 143–170. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.69.14bos>
- Bossong, Georg. 1998. Le marquage différentiel de l'object dans les langues de l'Europe. In Jack Feuillet (ed.), *Actance et valence dans les langues d'Europe*, 193–259. Berlin: Mouton de Gruyter.
- Cornilescu, Alexandra. 2000. Notes on the interpretation of the prepositional accusative in Romanian. *Bucharest Working Papers in Linguistics* 2(1). 91–106.
- Cornilescu, Alexandra & Alina Tigău. 2017. *On the competition between differentially marked objects and indirect objects in Romanian*. Paper presented at Going Romance 2017 University of Bucharest. 2017.
- D'Alessandro, Roberta & Ian Roberts. 2010. Past participle agreement in Abruzzese: Split-auxiliary selection and the null subject parameter. *Natural Language & Linguistic Theory* 28(1). 41–72. <https://doi.org/10.1007/s11049-009-9085-1>
- Davis, Henry. 2018. 'Nominalization' as predicativization in Lillooet and the nominal mapping parameter. Paper presented at GLOW 2018, Budapest, April 10–14.
- Espinal, Maria Teresa & Sonia Cyrino. 2017. The definite article in Romance expletives and long weak definites. *Glossa: A Journal of General Linguistics* 2(1). 23. 1–26. <https://doi.org/10.5334/gjgl.160>
- Giusti, Giuliana. 1993. *La sintassi dei determinanti*. Padua: Unipress.
- Harley, Heidi & Elizabeth Ritter. 2002. Person and number in pronouns: A feature-geometric approach. *Language* 78(3). 482–526. <https://doi.org/10.1353/lan.2002.0158>
- Kalin, Laura. 2018. Licensing and differential object marking. The view from Neo-Aramaic. *Syntax* 21(2). 112–159. <https://doi.org/10.1111/synt.12153>
- Ledgeway, Adam. 2000. *A comparative syntax of the dialects of southern Italy: A minimalist approach*. Oxford: Blackwell.
- Ledgeway, Adam, Norma Schifano & Giuseppina Silvestri. 2019. Differential object marking and the properties of D in the dialects of extreme south of Italy. *Glossa: A Journal of General Linguistics* 4(1). 1–25. <https://doi.org/10.5334/gjgl.569>

- Levin, Theodore. 2015. Licensing without case. Doctoral dissertation, MIT.
- Longobardi, Giuseppe. 2008. Reference to individuals, person, and the variety of mapping parameters. In Henrik Høeg Müller & Alex Klinge (eds.), *Essays on nominal determination. From morphology to discourse management*, 189–213. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.99.11lon>
- Longobardi, Giuseppe & Cristina Guardiano. 2009. Evidence for syntax as a signal of language relatedness. *Lingua* 119(11). 1679–1706. <https://doi.org/10.1016/j.lingua.2008.09.012>
- López, Luis. 2012. *Indefinite objects: Scrambling, choice functions and differential marking*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9165.001.0001>
- Loporcaro, Michele. 1998. *Sintassi comparata dell'accordo participiale romanzo*. Turin: Rosenberg & Sellier.
- Loporcaro, Michele. 2010. The logic of past participle agreement. In Roberta D'Alessandro, Adam Ledgeway & Ian Roberts (eds.), *Syntactic variation: The dialects of Italy*, 225–243. Cambridge: Cambridge University Press.
- Manzini, M. Rita & Ludovico Franco. 2016. Goal and DOM datives. *Natural Language & Linguistic Theory* 34(1). 197–240. <https://doi.org/10.1007/s11049-015-9303-y>
- Miyagawa, Shigeru. 2010. *Why agree? Why move?* Cambridge, MA: MIT Press.
- Miyagawa, Shigeru. 2017. *Agreement beyond phi*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/10958.001.0001>
- Moulton, Keir. 2013. Raising from the dead. *Linguistic Inquiry* 44(1). 157–167. https://doi.org/10.1162/LING_a_00123
- Nevins, Andrew. 2007. The representation of third person and its consequences for person-case effects. *Natural Language & Linguistic Theory* 25(2). 273–313. <https://doi.org/10.1007/s11049-006-9017-2>
- Ormazabal, Javier & Juan Romero. 2013a. 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>
- Ormazabal, Javier & Juan Romero. 2013b. Non accusative objects. *Catalan Journal of Linguistics* 12. 155–173. <https://doi.org/10.5565/rev/catjl.65>
- Richards, Marc. 2008. Defective agree: Case alternations, and the prominence of person. In Marc Richards & Andrej L. Machukov (eds.), *Scales*, *Linguistische Arbeits Berichte* 86, 137–161. Leipzig: Universität Leipzig.
- Rodríguez-Mondoñedo, Miguel. 2007. The syntax of objects. Agree and differential object marking. Doctoral dissertation, University of Connecticut.
- Suñer, Margarita. 1988. The role of agreement in clitic-doubled constructions. *Natural Language & Linguistic Theory* 6(3). 391–434. <https://doi.org/10.1007/BF00133904>
- Tigău, Alina. 2011. *Syntax and interpretation of the direct object in Romance and Germanic languages with an emphasis on Romanian, German, Dutch and English*. Bucharest: Editura Universității din București.
- Torrego, Esther. 1998. *The dependencies of objects*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/2337.001.0001>
- Williams, Edwin. 1983. Against small clauses. *Linguistic Inquiry* 5(2). 225–243.
- Wiltschko, Martina. 2014. *The universal structure of categories: Towards a functional typology*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9781139833899>
- Zamparelli, Roberto. 2000. *Layers in the determiner phrase*. New York: Garland Publishing (Routledge).

The effects of language ecology on syntactic structure

A look at Kristang and Makista

Robert W. Laub

SOAS University of London

Kristang and Makista are two related Portuguese-lexified creoles of southeast Asia. While they are generally mutually intelligible, there are differences between them. A look at the contact languages found in Malacca and Macau can give us ideas as to where these differences come from. Building on Ansaldo's notion of language ecology, this study will look at the histories of Kristang and Makista, and the features that exhibit these differences. These histories will be taken into account when looking at the differences, and evidence suggesting the connection between language ecology and morphosyntactic structures will be presented. The two features discussed are genitive patterns, and cleft constructions.

Keywords: creole, language ecology, Makista, Kristang, substrate

1. Introduction

The present study investigates two related creoles in southeast Asia: Kristang of Malacca, Malaysia, and Makista of Macau. The sociohistorical backgrounds of these two cities will be investigated as to how they influenced the languages' structures. Important factors to consider in this study range from European colonialism to present-day demographics and language policies. Two features are investigated in this chapter: genitives and cleft constructions. The genitive features in the two languages are near identical, so a quantitative method has been applied. I have used a qualitative method for discussing the cleft constructions, as they have significantly different structures.

2. Data sources

The Kristang data sources in this contribution are from three sources: the writings of Joan Marbeck (1995), a Kristang speaker from Malacca and example sentences in Alan Baxter's 1988 grammar of Kristang. Makista data is much harder to access, so the sources in this chapter are much more limited. The prose writings of José dos Santos Ferreira, a native speaker of Makista, have been used as data for the purposes of this study.

3. Competing views in creole studies

Since individual creoles were identified as languages separate from their lexifiers, linguists have pointed out similarities between some of them and developed theories of creole development and potential distinctiveness. One such theory that argues in favor of creoles' distinctiveness as a class of languages is Bickerton's Language Bioprogram Hypothesis (LBH) (1984). He argues that the structures of contact varieties are determined by the internal language organ, and there is a resetting of parameters as a creole emerges from a pidgin. In this hypothesis, the super- and substrate languages play a little role in the syntactic structure of the creoles.

Bakker et al. (2011) further argued this case, in that creole languages are typologically distinct. Looking at features of creoles compared to non-creoles, they showed that creoles clustered together. Adding to similar claims by McWhorter (2001, 2005) and Parkvall (2008), they used a quantitative method to look at features deemed prototypical of creoles (Holm & Patrick 2007). Looking at creoles from a variety of lexifiers, geographical regions, and types, they argued that regardless of the creole's origin, it will have features that are deemed typical of this language group.

Youssef (1988, 1990), however, argued that Bickerton presumed that pidgins were formed in a vacuum, ignoring the fact that people acquiring creoles will have input from a variety of languages.

DeGraff (2003, 2005) argued that Eurocentric viewpoints led to the view that creoles are distinct. This goes back to views that creole varieties were viewed as inferior versions of the superstrates. Features and syntactic history can be found in both languages considered creoles and those not. One such Example (DeGraff 2009) is the negation markers in Haitian and English. In both languages, the negation marker precedes the verb, but both Haitian and English have ancestor languages with a negation after the verb (French and Old English, respectively). Despite this similarity, Haitian is considered a creole while English is not.

Ansaldo (2009) noted that environments in which languages are used affect languages' development. Social aspects in language development and evolution cannot be ignored, especially in a multicultural area such as southeast Asia, as the geography of the area was conducive to intra-Asian trade among, for example, Chinese, Indians, and Arabs.

4. Historical background

Lying on quickest route between the Pacific and Indian Oceans, Malacca has long been an important trading port (Ansaldo 2009). Seasonal trade winds aided this (McPherson 1993), which further helped traders from far away reach the area. Asian traders were active in this region long before the Europeans arrived, and they started communities of their own within the region.

The language spoken by the locals in this period was Malay, and a contact variety, Bazaar Malay, emerged. More isolating than literary Malay, it was the lingua franca among Asian traders in the region (Ansaldo 2009).

After expanding a trading network around the African and Indian coasts, the Portuguese invaded Malacca in 1511. The Malacca they encountered was a cosmopolitan city in which Portuguese sailors married locals and had children with them. The language that emerged from this mixed community is Kristang (Ostler 2005).

Malacca became a center for Portuguese trade, and Portugal expanded its presence in the region as far as Japan, and as the route from Malacca was long, they created a stopover colony in China in 1557, which became Macau (Arana Ward 1977; Sit et al. 2012; Pinharanda Nunes 2013). After the invasion, China prohibited its own citizens from entering Macau, so the Portuguese settled it with people from elsewhere, including Malacca Eurasians, who brought with them their language which turned into Makista.

4.1 Kristang

The Portuguese controlled Malacca long enough to create a Eurasian community, but actual Portuguese control lasted only about a century. As the Dutch entered spice trade, they took control from Portugal. The Dutch lost control of Malacca to the British, who held onto their colony until independence in 1957 (Baxter 1988). Throughout this period, Kristang continued to be spoken, although contact with Portuguese was almost completely cut off during this time. Because of its location, the language has much influence from its substrate, Malay, and currently speakers of Kristang are shifting to English, the lingua franca of Malaysia. Most speakers

of Kristang are in Malacca, but some live elsewhere in Malaysia and in Singapore (Baxter 2005). Kristang has a few thousand speakers, and members of the community are active in maintaining and revitalizing the language.

4.2 Makista

Makista has its origins with Kristang speakers who settled in Macau after the Portuguese takeover. Portugal held onto Macau until 1999, and these centuries of control have led to Portuguese remaining as an official language in Macau, currently a Special Administrative Region of China (Cheng 1999). This distinction means that the Portuguese language was more present in Macau, and resulted in decreolization of Makista (Charpentier 1995; Pinharana Nunes 2012, 2013). While Makista has a Malayo-Portuguese origin and retains many of the Malay-influenced features that are present in Kristang, the main contact languages of its speakers have been Cantonese and Portuguese, and English via Hong Kong. This language is critically endangered, with fewer than 50 native speakers. There are efforts to expand the usage of Makista (Lewis et al. 2016).

4.3 Kristang and Makista

Previous studies in a similar vein include Holm & Patrick (2007), who wrote a volume comparing the syntax of various creole languages. Other works have looked at Kristang (Baxter 1988) and Makista (Arana-Ward 1977; Ansaldo & Matthews 2004; Pinharanda Nunes & Baxter 2004; Pinharanda Nunes 2008, 2012; Avram 2015; Arcodia 2017; Lebel 2018). There have also been comparisons of Luso-Asian creoles from broader perspectives (Cardoso et al. 2012). As Kristang and Makista are related creoles, this adds significance to the present study as it can show what happens when languages develop divergently from a lexifier, and the factors behind the divergence.

5. Common features

Both languages have similar Malay-influenced features. One of the most prominent is the usage of preverbal particles to mark aspect on a verb, contrasting with the Portuguese feature of using verbal inflection.

- (1) a. Malay (native speaker)
Saya sudah makan pisang.
 1SG PFV eat banana
 ‘I’ve (already) eaten the banana.’

- b. Kristang (Thurgood & Thurgood 1996)
Já susude ungua témpu.
 PFV occur one time
 “It happened one time.”
- c. Makista (Santos Ferreira 1967)
Queléora iou já fazê sai Macau Sã Assi.
 when 1SG PFV make go out Macau Sã Assi
 “When I published *Macau Sã Assi*” (Santos Ferreira 1967)

6. Genitive

Baxter & Bastos (2012) identified two forms of genitive in Kristang: pre- and post-nominal. The prenominal genitive is syntactically parallel to the Portuguese, and consists of the possessee, followed by *of* and the possessor.

- (2) a. Portuguese
Palácio do rei
 Palace of the king
 “Palace of the king”
- b. Kristang (Baxter & Bastos 2012: 67)
Palasu di re
 Palace of king
 “Palace of the king”

The postnominal genitive consists of the possessor followed by a genitive particle, and then the possessee. This has parallels with many southeast and south Asian languages, and Baxter & Bastos speculate that the feature originated in south Asia (see Example (3) below) and was later reinforced as contact varieties were brought to Malay-speaking areas.

- (3) a. Marathi (Clements 1996: 140)
Kapil tsa gher
 Kapil GEN house
 “Kapil’s house”
- b. Bazaar Malay (Baxter 1988: 92)
Gua punya rumah
 1SG GEN house
 “My house”
- c. Kristang (Baxter & Bastos 2012: 60)
Singapura sa jenti lo beng Malaka
 Singapore GEN people FUT IRR come Malacca
 “People from Singapore will come to Malacca”

Both of these features are found in Kristang, but different semantic categories tend to trigger one pattern or another. In their study, Baxter & Bastos looked at tokens of genitive and assigned them to the following categories: kinship, body parts, ownership, other interpersonal relationships, classificatory, part-whole, spatial/locative, origin/source, and material composition.

Table 1 is adapted from Baxter & Bastos (2012: 66) and shows the number and percentage of the types of genitive in each semantic category

Table 1. Kristang genitive patterns (Baxter & Bastos 2012)

Semantic category	Number of postnominal	% of postnominal per category	Number of prenominal	% of postnominal per category
Kinship	500	100%	0	0%
Body part	80	96%	3	4%
Ownership	250	94%	16	6%
Other relationship	40	89%	5	11%
Classificatory	6	50%	6	50%
Part-whole	8	42%	11	58%
Spatial/locative	10	16%	51	84%
Origin/source	5	13%	34	87%
Material composition	4	6%	62	94%

As shown, the postnominal genitive is highly preferred for terms regarding human relations and inalienable possessions, whereas non-human, and especially inanimate, possessors prefer the prenominal type.

Makista has the same two features.

- (4) Makista (Santos Ferreira 1973)
Quarto di capitám
 room of captain
 “The captain’s room”
- (5) Makista (Santos Ferreira 1973)
Nhu-nhúm- sa siara
 men GEN wife
 “Men’s wives”
- (6) Makista (Santos Ferreira 1973)
Iou- sua coraçam. Ta querê pará!
 1SG GEN heart PROG want stop
 “My heart. It’s stopping!”

(7) Makista

(Santos Ferreira 1973)

Dessá olá iou- sua passaporte sã pa basso.

Leave see 1SG GEN passport COP to bottom

“See to it that my passport is put away.”

For the purposes of the present chapter, I looked at instances of genitive in Santos Ferreira’s 1967 volume *Macau Sã Assi*.

Table 2 is adapted from Baxter and Bastos with the following changes: as there were no instances of other interpersonal relationships, that category has been omitted; also, I have expanded the spatial/locative to include temporal.

Table 2. Makista genitive patterns

Semantic category	Number of postnominal	% of postnominal per category	Number of prenominal	% of postnominal per category
Kinship	22	81%	5	19%
Body part	11	79%	3	21%
Ownership	17	59%	12	41%
Classificatory	0	0%	4	100%
Part-whole	0	0%	7	100%
Spatial/locative/ temporal	0	0%	12	100%
Origin/source	0	0%	20	100%
Material composition	0	0%	4	100%

While prenominal genitive was found to be used in nearly all the categories, and with various possessors, the postnominal was found only in the kinship, body parts and ownership categories. All of the possessors were human and most were found with pronouns.

With comparison to Kristang patterns in Table 1, this shows an expansion of prenominal usage, even going into the kinship category (absent in the Kristang corpus). Portuguese uses a feature identical to the prenominal pattern, and continual contact with Portuguese could have eroded away at the postnominal genitive.

7. Cleft constructions

Cleft constructions in Kristang are formed using the structure *teng* (copula) + definite NP + relative clause (Baxter 1988).

- (6) Kristang (Baxter 1988)
Teng bos sa familia ki yo ta lantah, retu?
 COP 2SG GEN child REL 1SG PROG carry right
 “It’s your child I am carrying, right?”

Similar structures can be found in both Portuguese and Malay.

- (7) a. Portuguese (native speaker)
É a sua criança que estou levando, não é?
 COP your child REL am carrying NEG COP
 “It’s your child I am carrying, right?”
- b. Malay (native speaker)
Kopi yang anda sedang minum, betul kah?
 coffee REL 2SG PROG drink right Q
 “It’s coffee you’re drinking, right?”

Makista exhibits a cleft construction not found in Kristang or in its contact languages. This is the usage of the copula (*sã* in Makista), which likely has its origins in Cantonese, as Sinitic languages exhibit a similar structure.

- (8) a. Cantonese (Matthews & Yip 1994)
Ngóh haih duhk Yíngmán ge
 1SG COP study English PRT
 “It’s English I study.”
- b. Makista (Santos Ferreira 1967)
Iou sã já nacê na Macau
 1SG COP PFV born in Macau
 “It’s that I was born in Macau.”

This Sinitic-style copular cleft is not possible in Kristang. Its close resemblance to Sinitic suggests that it was an innovation that occurred in Macau, rather than a feature found in both creoles that was lost in Kristang.

8. Conclusion

This paper has investigated two related creoles, which is significant as it gives an opportunity to look at the types and motivations of divergence among languages. Both Kristang and Makista emerged in situations that were ripe for cross-linguistic contact, and contact involving very different languages. While many creole studies (Bickerton 1981, 1984; Bakker et al. 2011; Bakker 2014) tend to focus on similarities between creoles, the differences are another important factor in their syntactic analysis. Another issue is the choice of languages being compared in these studies. While Bakker, Holm, and others take data from dozens of languages around the world, this study zooms into two languages that have a single common ancestor. This means that features in common are less likely to be coincidental, and helps us look at the morphosyntactic differences from a sociohistorical viewpoint.

Ansaldo's notion of language ecology points us in the direction of substrates and other contact languages within a creole's setting. As languages interacted with others and new varieties emerged, they took from the languages around them the features that they currently use. The goal of this study was not to show that creoles as a typological category do not exist, but to provide evidence of substrate and contact influence that might otherwise be discarded. More features will need to be looked at in the future, although this is difficult given the paucity of the available data. However, such an endeavor should take into account the relationship between two languages, especially creoles, before embarking on a comparative study.

Abbreviations used

1SG	1st person singular pronoun	GEN	genitive marker
2SG	2nd person singular pronoun	PFV	perfective marker
3SG	3rd person singular pronoun	PROG	progressive marker
COP	copula	PRT	particle
FUT IRR	future/irrealis marker	REL	relative pronoun
FP	final particle		

References

- Ansaldo, Umberto. 2009. *Contact languages: Ecology and evolution in Asia*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511642203>
- Ansaldo, Umberto & Stephen Matthews. 2004. The origins of Macanese reduplication. In Geneviève Esquire & Armin Schwegler (eds.), *Creoles, Contact and language change: Linguistic and social implications*, 1–19. Amsterdam: John Benjamins. <https://doi.org/10.1075/cll.27.02ans>

- Arana-Ward, Marie. 1977. A synchronic and diachronic investigation of Macanese: the Portuguese-based creole of Macau. MA thesis, University of Hong Kong.
- Arcodia, Giorgio. 2017. On Sinitic influence on Macanese. *Language Ecology* 1(2). 158–184. <https://doi.org/10.1075/le.1.2.03arc>
- Avram, Andrei. 2015. Reduplication in Macao Creole Portuguese and its origins. *Études créoles* 33(2). 27–51.
- Bakker, Peter. 2014. Creoles and typology: Problems of sampling and definition. *Journal of Pidgin and Creole Languages* 29(2). 437–455. <https://doi.org/10.1075/jpcl.29.2.09bak>
- Bakker, Peter, Aymeric Daval-Markussen, Mikael Parkvall & Ingo Plag. 2011. Creoles are typologically distinct from non-creoles. *Journal of Pidgin and Creole Languages* 26(1). 5–42. <https://doi.org/10.1075/jpcl.26.1.02bak>
- Baxter, Alan. 1988. *A grammar of Kristang (Malacca Creole Portuguese)*. Canberra: Australian National University.
- Baxter, Alan. 2005. Kristang (Malacca Creole Portuguese) – a long-time survivor seriously endangered. *Estudios de sociolingüística* 6(1). 1–37.
- Baxter, Alan & Augusta Bastos. 2012. A closer look at the post-nominal genitive in Asian Creole Portuguese. In Hugo Cardoso, Alan Baxter & Mário Pinharanda Nunes (eds.), *Ibero-Asian creoles: Comparative perspectives*, 47–49. Amsterdam: John Benjamins. <https://doi.org/10.1075/cll.46.03bax>
- Bickerton, Derek. 1981. *Roots of language*. New York: Karoma Publishers.
- Bickerton, Derek. 1984. The language bioprogram hypothesis. *Behavioral and Brain Sciences* 7. 173–188. <https://doi.org/10.1017/S0140525X00044149>
- Cardoso, Hugo, Alan Baxter & Mário Pinharanda Nunes (eds.). 2012. *Ibero-Asian creoles: Comparative perspectives*. Amsterdam: John Benjamins. <https://doi.org/10.1075/cll.46>
- Charpentier, Jean-Michel. 1995. O processo de descriolização no caso do crioulo macaísta de Macau a passagem dum basilecto para um acrolecto. *Papia* 4(1). 21–31.
- Cheng, Christina Miu Bing. 1999. *Macaí: A cultural Janus*. Hong Kong: Hong Kong University Press.
- Clements, J. Clancy. 1996. *The genesis of a language: The formation and development of Korlai Portuguese*. Amsterdam: John Benjamins. <https://doi.org/10.1075/cll.16>
- DeGraff, Michel. 2003. Against creole exceptionalism. *Language* 79(2). 391–410. <https://doi.org/10.1353/lan.2003.0114>
- DeGraff, Michel. 2005. Linguists' most dangerous myth: The fallacy of creole exceptionalism. *Language in Society* 34(4). 533–591. <https://doi.org/10.1017/S0047404505050207>
- DeGraff, Michel. 2009. Creole exceptionalism and the (mis)education of the creole speaker. In Anne Kleifgen & George Bond (eds.), *The languages of Africa and the diaspora: Educating for language awareness*, 124–144. Bristol: Multilingual Matters. <https://doi.org/10.21832/9781847691354-011>
- Holm, John & Peter Patrick (eds.). 2007. *Comparative creole syntax: Parallel outlines of 18 creole grammars*. London: Battlebridge.
- Lebel, Alexandre. 2018. Seven ways of saying *no*: The negation system in Macau Creole Portuguese. *Papia* 28(2). 157–200.
- Lewis, M. Paul, Gary F. Simons & Charles D. Fennig (eds.). 2016. *Ethnologue: Languages of the world* (19th edn.). Dallas: SIL.
- Marbeck, Joan Margaret. 1995. *Ungua Adanza*. Lisbon: Calouse Gulbenkian Foundation.
- Matthews, Stephen & Virginia Yip. 1994. *Cantonese: A comprehensive grammar*. New York: Routledge.

- McPherson, Kenneth. 1993. *The Indian Ocean: A history of people and the sea*. Oxford: Oxford University Press.
- McWhorter, John. 2001. The world's simplest grammars are creole grammars. *Linguistic Typology* 5(2/3). 125–166. <https://doi.org/10.1515/lity.2001.001>
- McWhorter, John. 2005. *Defining creole*. New York: Oxford University Press.
- Ostler, Nicholas. 2005. *Empires of the world: A language history of the world*. New York: Harper Collins.
- Parkvall, Mikael. 2008. The simplicity of creoles in a cross-linguistic perspective. In Matti Miestamo, Kaius Sinnemäki & Fred Karlsson (eds.), *Language complexity: typology, contact, change*, 265–286. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.94.17par>
- Pinharanda Nunes, Mário. 2008. Os demonstrativos em maquista: uma análise morfosintáctica contrastive. *Papia* 18. 7–21.
- Pinharanda Nunes, Mário. 2012. Traces of superstrate verb inflection in Makista and other Asian-Portuguese Creoles. In Hugo Cardoso, Alan Baxter & Mário Pinharanda Nunes (eds.), *Ibero-Asian creoles: Comparative perspectives*, 289–326. Amsterdam: John Benjamins. <https://doi.org/10.1075/cll.46.11pin>
- Pinharanda Nunes, Mário. 2013. Socio-historical factors in changes in the creole matrix of Makista. In George Wei & Katrine Wong (eds.), *Macau – Cultural interaction and literary representation*, 25–41. New York: Routledge.
- Pinharanda Nunes, Mário & Alan Baxter. 2004. Os marcadores pre-verbais no crioulo de base lexical portuguesa de Macau. *Papia* 14. 31–46.
- Santos Ferreira, José. 1967. *Macau sã assi*. Macau: Tipografia da Missão Padroado.
- Santos Ferreira, José. 1973. *Qui-nova chenko*. Macau: Tipografia da Missão Padroado.
- Sit, Victor F. S., Vivian Hui & Gladys Leung. 2012. *Macau through 500 years*. Singapore: Silkroad Press.
- Thurgood, Ela & Graham Thurgood. 1996. Aspect, tense, or aktionsart? The particle *ja* in Kristang (Malacca Creole Portuguese). *Journal of Pidgin and Creole Languages* 11(1). 45–70. <https://doi.org/10.1075/jpcl.11.1.04thu>
- Youssef, Valerie. 1988. The language bioprogram hypothesis revisited. *Journal of Child Language* 15. 451–458. <https://doi.org/10.1017/S0305000900012459>
- Youssef, Valerie. 1990. On the confirmation of bioprograms. *Journal of Child Language* 17. 233–235. <https://doi.org/10.1017/S0305000900013209>

The syntactic distribution of *raddoppiamento fonosintattico* in Cosentino

A phase-theoretic account

Adam Ledgeway

University of Cambridge

This chapter undertakes an overview of the structural conditions regulating the distribution of a phonological fortition process, *raddoppiamento fonosintattico* “phonosyntactic doubling”, in the Calabrian dialect of Cosenza, focusing on some particularly striking cases which also reveal the role of phases in constraining the application of RF. On the one hand, the data highlight the advantages of interpreting locality not just narrowly in terms of the three core structural configurations Spec-Head, Head-Head and Head-Comp, but also more broadly in terms of phasal domains, showing how different phonological realizations represent the spell-out of deep syntactic differences mapped at the interface between narrow syntax and PF (Phonological Form). On the other, the theoretical assumptions assumed here provide us with the key to understanding some intriguing empirical generalizations about the role of RF in signalling, among other things, informational structure content and binding relations, which, in turn, throw new light on current theoretical assumptions about clause structure and the nature of phases.

Keywords: syntax-phonology interface, phase, locality, *raddoppiamento fonosintattico*, information structure

1. Introduction

It is well known that many phonological processes, in particular external sandhi phenomena such as Welsh soft mutation (cf. Roberts 2005), show sensitivity to syntactic information, insofar as their surface distribution can only be understood by making reference to various structural constraints in the mapping from syntax-to-phonology (cf. Selkirk 1984, 1986; Kaisse 1985; Selkirk & Tateischi 1988; Inkelas & Zec 1990; Truckenbrodt 1999). Romance too has been shown to present various such cases (for an overview see Sampson 2016: § 40.3) including, among

others, French *liaison* (Selkirk 1974; Morin & Kaye 1982; Durand & Lyche 2008; Bonami et al. 2014; Masutti 2016), intonational phrasing (Elordieta et al. 2003; D’Imperio et al. 2005; Rao 2008) and various vocalic processes in the dialects of southern Italy (Savoia 1987, 2015: Chapter 6; Rizzi & Savoia 1993; Silvestri 2009; Manzini & Savoia 2016). In what follows we consider how in the southern Italian dialect of Cosenza (Calabria) syntactic constraints determine the distribution of *rad-doppiamento fonosintattico* (RF) ‘phonosyntactic doubling’, an initial-consonant lengthening process which originates as an external sandhi assimilation triggered by a small class of words that historically ended in a final consonant such as Cosentino *cchiù* (< PLUS) ‘more’ (cf. (1a)).¹ While superficial phonological adjacency appears sufficient for RF to apply, it has been known since at least D’Ovidio (1874: 179) that the process in southern Italy is also constrained by structural conditions, a situation insightfully captured by Fanciullo’s (1986: 88) observation that RF ‘occurs only if the lexeme which causes it constitutes, together with the item it acts on, a *minimal phrase* – a kind of hierarchically superior word’ (cf. also Andalò 1991; Loporcaro 1997: 49; De Blasi & Imperatore 2000: 49–50; Savoia 2015: 436–441). In short, Fanciullo’s intuition is to link the phonological licensing of RF to constituency, hence its absence in (1b), where post-adjectival *cchiù* continues to modify, and hence form an immediate constituent with, the adjective *granne*, but not the linearly adjacent prepositional complementizer (DE > *di*) ‘i’ introducing the standard of comparison.²

- (1) a. *Cchiù* + *granne* > *cchiù ggranne*
 more tall more tall
 ‘Taller’
 b. *Granne cchiù* ‘i / **ddi* *mia*
 tall more of of me
 ‘Taller than me’

Developing this idea further, Ledgeway (2009a: 46–47) shows how in Neapolitan the relevant locality constraint on the application of RF encapsulated in Fanciullo’s

1. Cf. Rohlfs (1966: 235–238); Loporcaro (1988, 1997); Maiden (1995: 72–76); Fanciullo (1997); Sampson (2016: 675–676). For expository convenience, consonantal lengthening is indicated orthographically with an initial double consonant in bold-type, while acknowledging that the relevant contrast is not simply one of length but, rather, one of fortition involving in some cases a change of manner and of place of articulation or the restoration of an underlying consonant (cf. Savoia 2015: 415–416; Ledgeway 2016: 252–253), e.g., *cchiù* ‘more’ + *jancu* ‘white’ > *cchiù gghjancu* ‘whiter’ (/j/ > [j:]); *cchiù* ‘more’ + *a* ‘the’ + *carne* ‘meat’ > *ccchiù ra carne* ‘plus the meat’ (/∅/ > [r / r]).

2. Unless otherwise indicated, all examples are from the dialect of Cosenza.

‘minimal phrase’ can be broken down into the three core structural configurations of Spec-Head (2a), Head-Head (2b) and Head-Comp(lement) (2c), with RF failing to apply in linearizations falling outside of these configurations such as (3a–b).

- (2) a. *Accussi ffriddo* (Neapolitan; Spec-Head)
 so cold
 “So cold”
- b. *’O ssaccio.* (Neapolitan; Head-Head)
 it= know.PRS.1SG
 “I know it/that.”
- c. *So’ ppapà.* (Neapolitan; Head-Comp)
 be.PRS.1SG dad
 “I am a dad.”
- (3) a. *Accussi *(c)chiammaie sùbbeto.* (Neapolitan)
 thus call.PST.3SG immediately
 “Therefore he called at once.”
- b. *So’ *(c)comme rice tu.* (Neapolitan)
 be.PRS.3PL as say.PRS.2SG you
 “They are precisely as you say.”

However, such structural conditions are not sufficient to capture the distribution of RF in all southern varieties, as revealed by Biberauer & D’Alessandro’s (2006) analysis of the eastern Abruzzese dialect of Arielli where the application of RF in otherwise identical Head-Comp linearizations with passive/active auxiliary BE (e.g., SUM > *so* “be.PRS.1SG”) and perfective participle (e.g., *viste* “seen”) in (4) is only licensed under the passive interpretation, but not its active variant.

- (4) *So vviste / viste.* (Ariellese)
 be.PRS.1SG seen seen
 “I am seen / I have seen.”

Adopting a cyclical approach to Spell-Out, Biberauer & D’Alessandro successfully derive this contrast from the variable phasal status of the participial *vP* which, on standard assumptions, only constitutes a phase when active (see also D’Alessandro & Scheer 2015: 610–614). Consequently, in the active Head-Comp sequence consonantal lengthening fails to obtain because active auxiliary and participle are sent to PF in separate cycles, whereas in the passive string both passive auxiliary and participle are contained within the same higher CP phase and sent to PF together in the same cycle where the auxiliary can license consonantal lengthening of the adjacent passive participle. In a similar vein, Silvestri (2007, 2014) shows how in the northern Calabrian dialect of Verbicaro the distribution of RF is also restricted by phasal domains.

In the wake of these latter studies, this chapter undertakes a broad overview of the structural conditions regulating Cosentino RF, focusing on some particularly striking cases which also reveal the role of phases in constraining the application of RF. On the one hand, the data highlight the advantages of interpreting locality not just narrowly in terms of the three core configurations noted above in Ledgeway (2009a), but also more broadly in terms of phasal domains, showing how different phonological realizations represent the spell-out of deep syntactic differences mapped at the interface between narrow syntax and PF. On the other, the theoretical assumptions assumed here provide us with the key to understanding some intriguing empirical generalizations about the role of RF in signalling, among other things, informational structure content and binding relations, which, in turn, throw new light on current theoretical assumptions about clause structure and the nature of phases.

2. Cosentino RF: Initial overview

2.1 RF-triggers

Below in (5a–h) I provide an exhaustive list of RF-triggers in Cosentino:

- (5) a. Conjunctions: (ET >) *e* “and” (*e ddumani* “and tomorrow”), (AC >) *a* “and” (limited to now lexicalized DECEM “ten” + AC “and” + SEPTEM/NOUEM “seven/nine” > *diciassette/diciannove* “seventeen/eighteen”)
- b. Quantifiers: (OMNES >) *ogni* “each” (*ogni ccappieddu* “each hat”), (TRES >) *tri* “three” (*tri ffimmine* “three women”);
- c. *Wh*-phrases: (QUID >) *cchi* “what?!/” (*Cchi cciuotu?!/* “What (an) idiot?!/!”)
- d. (Religious) titles: (MATER >) *matre* “mother” (*Matre Tteresa* “Mother Theresa”), (PATER >) *patre* “father” (*Patre Ppio* “Father Pio”)
- e. Negators: (NON >) *no/u(n)* “not” (*no ssenza tia* “not without you”), (NEC >) *né* “neither, nor” (*né Rrosina* “nor Rosina”)
- f. Prepositions: (AD >) *a* “to, at” (*a Mmilanu* “to/in Milan”), (CUM >) *ccu* “with” (*ccu mmia* “with me”), (PER >) *ppi* “for” (*ppi bbua* “for you”)
- g. Adverbs: (*ACCU-HAC >) *ccà* “here” (*ccà ssutta* “down here [*lit.* here down]”), (ILLAC >) *ddà* “there” (*ddà bbicinu* “near there [*lit.* there near]”), (PLUS >) *cchiù* “more” (*cchiù ttiempu* “more time”), (*ACCU-SIC >) *accussi* “so, thus” (*accussi fforte* “so strong”)
- h. Verbs: (SIS >) *si* “be.PRS.2SG” (*si ppicceriddu* “you are small”), (SUN(T) >) *su* “be.PRS.3PL” (*su ppicceriddi* “they are small”), and all 3SG finite verb forms (< -T; e.g., *frica ttuttu* “he steals everything”)

Although the inventory proves extremely limited, with RF licensed after just a handful of members of individual word classes (cf. *e* (< ET) *ddumani* “and tomorrow” vs

o (< AUT) (**d*)*dumani* “or tomorrow”),³ the incidence of RF is anything but rare, in that the relevant words represent some of the most frequently recurring grammatical items in the dialect. In the case of (5h) the effects of original underlying final -T in all 3SG finite verb forms guarantees that the actual number and frequency of lexical items able to trigger RF are hardly negligible (Loporcaro 1997: 114–117).⁴ Moreover, RF has been analogically extended to the 1SG in paradigms where 1SG and 3SG are today homophonous (cf. CLAMABA-M/-T > *chiamava* “I/(s)he called”, CLAMAUISSE-M/-T > *chiamassa* “I/(s)he called/would call”, and CLAMAUERA-M/-T > *chiam(e)ra* “I/(s)he would call”), thereby considerably increasing further the distribution and incidence of RF in the dialect, e.g., *spinnissa ssordi* “I/(s)he would spend money”.⁵

2.2 Distribution: Core syntactic configurations

In terms of the structural domains in which RF is licensed, as an initial generalization we note that it is readily licensed in all three of the core local configurations illustrated in (6a–c).⁶

- (6) a. [AdvP [Spec *Cchiù*] *cchianu*]
 more slowly
 “More slowly” (Spec-Head)
- b. [TP [T' [T° [Neg *Un*] [T [C[*tti*] [T *capisciu.*]]]]]
 not = you.SG.ACC= understand.PRS.1SG
 “I don’t understand you.” (Head-Head)
- c. [&P [Spec *Maria*] [&' *e* [DP *nnua*]]]
 Maria and we
 “Maria and us” (Head-Comp)

3. The small number of RF-triggers undoubtedly represents the cumulative effect of diachronic erosion of a once much bigger inventory – witness, for example, the RF effects of (PER+QUID >) *picchi* “why?”, (quid >) *chi* “that (complementizer)” and (*SED >) *si* “if” in neighbouring dialects (Rohlf 1966: 238) – progressively reduced on account of the lexically idiosyncratic and largely opaque distribution of RF in synchronic terms (Loporcaro 1988; Ledgeway 2009a: 39–48).

4. Note that, contrary to what is claimed in Loporcaro (1998), in pre-pausal position the original final dental (viz. -*de*) resurfaces also in the urban dialect of Cosenza, see contrasts such as *Cicciu scrivìa nna littera/scriviade* “Cicciu was writing a letter/was writing.”

5. It is highly unlikely that RF in the 1SG represents the original assimilation of final -*m*, since, with the exception of monosyllables (cf. CUM > *ccu* [+RF] “with”, as well as Fr. *rien* “nothing” < ACC.SG REM “thing”), the latter is known to have weakened very early in the history of Latin.

6. For expository clarity, many of the following structural representations have been simplified, omitting, for example, lower copies and all projections not immediately relevant to the discussion.

In (6a) the degree adverb *cchiù* enters into a Spec-Head relation with the manner adverb *chianu* lexicalizing its specifier, a sufficiently local position from which to license RF on the latter. In (6b) the object clitic *ti* and the clitic negator *un* both raise to left-adjoin into the finite verb (cf. Kayne 1991),⁷ itself raised to a functional head within the T-domain, placing the negator and clitic in a Head-Head relation as part of a complex head in which the former triggers RF on the latter. Finally, in (6c) the coordinating conjunction selects a pronominal complement, a D head with null NP complement, producing a canonical Head-Complement configuration in which the head triggers initial-consonantal lengthening of its pronominal complement.

However, these three configurations must also be understood more broadly to accommodate structures like (7a–c) in which doubler and doublee (henceforth Word₁ and Word₂), although linearly contiguous, are not immediately adjacent in structural terms.

- (7) a. [DP [Spec *Cchiù*] [D' D° __ [NP *ccasu*]]]
 more cheese
 “More cheese” (Spec-Head)
- b. [ForceP [Force' *Fa-* [FinP [Fin' [Fin° [Cl *llu*] [Fin° *fa* [v-VP *lu fa!*]]]]]]]]]
 do.IMP.2SG =it.ACC.MSG
 “Do it!” (Head-Head)
- c. [&P [Spec *Pane*] [& e [DP [D' D° __ [NP *ccasu*]]]]]
 bread and cheese
 “Bread and cheese” (Head-Comp)

In (7a) *cchiù* is now merged in Spec,DP, but the head of the projection remains empty leaving the adverb to modify the empty head's NP complement instantiated by the non-projecting nominal head *casu*. We thus obtain a Spec-Head configuration, albeit one in which between the adverb and the nominal head there intervenes a non-lexicalized D°. Similarly, in (7b) we adopt the idea that imperatival clauses fail to project any T-related functional structure,⁸ with the imperatival verb and the clitic *lu* raising to a low head in the C-domain (viz. Fin°),⁹ from which the verb subsequently excorporates to reach a higher C-related head (viz. Force°), where it can presumably license the imperatival illocutionary force of the clause (cf. Munaro 2004, 2010). Consequently, the verb and the now enclitic object pronoun come to

7. There are, of course, alternative analyses which posit that under proclisis such clitic elements lexicalize separate, albeit, adjacent heads above the verb which left-adjoins to a functional head within the T-domain (cf. Kayne 1994; Ledgeway & Lombardi 2005).

8. Cf. Rivero (1994a; b); Graffi (1996); Zanuttini (1997); Tortora (2014: Chapter 3 § 6); Ledgeway (2020).

9. Cf. Rivero (1994b); Rivero & Terzi (1995); Manzini & Savoia (2005, III: 388).

occupy adjacent but distinct head positions. Finally, in (7c) the complement of the coordinating conjunction is a bare nominal, in structural terms an NP embedded within a DP with a null head, such that between head and complement there intervenes a functional projection. In all three cases strict structural adjacency between Word₁ and Word₂ fails to obtain, but the configurations they engender are nonetheless sufficiently local to license RF in each case.

2.3 Phase-theoretic interpretation of locality

Evidence like that considered in (7a–c) forces us then to rethink the relevant locality conditions on the distribution of RF which appear to lend themselves more readily to an interpretation in terms of phase theory, rather than the more restrictive predictions of the three traditional core configurations in (6a–c). Indeed, examples like (7a–c) support the idea that the only relevant locality restriction on the adjacency requirement of RF is that Word₁ and Word₂ both surface in the same phasal domain. This is further demonstrated by the examples in (8a–c) which highlight how, although Word₁ and Word₂ are not immediately adjacent in structural terms, the licensing of RF is again ensured by their co-occurrence within the same phase at PF.

- (8) a. (*Un ssacciu*) [FocP [Spec *Cchi*] [Foc' Foc° ___ [FinP [TP... [AspP
not know.PRS.1SG what
ccunta Cciciu (?)]]]] (Spec...Head)
say.PRS.3SG Ciciu
“What is Ciciu saying? (I don’t know what Ciciu is saying.)”
- b. [ModP *Pò*... [AspP *pparta* [_v-VP *parta*.]]] (Head...Head)
can.PRS.3SG leave.INF
“He can leave.”
- c. [ModP *Vò*... [AspP [Spec *ttuttu*.]]] (Head...Comp)
want.PRS.3SG everything
“He wants everything.”

(8a) illustrates a Spec...Head configuration where Word₁ lexicalizes a specifier position and Word₂ a lower head position, between which there intervene several projections, even in non-cartographic analyses. In particular, the *wh*-phrase raises to Spec,FocP, but the verbal head only raises to a head within Cinque’s (1999) aspectual field above *v*-VP (Ledgeway & Lombardi 2005).

Similarly, in the Head...Head configuration in (8b) Word₁ *pò* lexicalizes a modal head in the higher portion of Cinque’s (1999) articulated clause structure, whereas the lexical infinitive instantiated by Word₂ raises to a clause-medial position (Ledgeway & Lombardi 2005: § 3.2) which, although at some structural distance from the former (cf. interpolation of adverbs such as *Pò ccertu/fforse/ssempe/*

ggjà parta “He can certainly/perhaps/always/already leave”), still occurs in the same phase as *pò*.

Finally, (8c) illustrates a Head...Comp configuration in which the modal *vò*, once again merged in Cinque’s higher modal field, finds itself at some considerable distance from its complement, the bare quantifier *tuttu* which lexicalizes the specifier of a plural completive aspectual projection situated in the lower portion of the aspectual field. Furthermore, this latter example highlights how RF does not necessarily have to apply to the head of the complement (here not lexicalized), but simply applies to the closest adjacent item contained within the complement, its specifier in this case. This explains why in the following sequences we can find RF on the head of the complement (9a), on its specifier (9b) or on a postnominal adjunct (9c). As a consequence, we can also find both Spec-Spec (10a) and Spec...Spec configurations (10b) where, as in the previous examples, Word₁ and Word₂ do not necessarily form a constituent (pace Fanciullo 1986), but RF signals their linear adjacency within the same phase.

- (9) a. [NumP *Tri* [NP *ggatti*]]
 three cats
 “Three cats”
- b. [NumP *Tri* [NP [Spec *ppoveri*] *gatti*]]
 three poor cats
 “Three poor cats”
- c. *Ni tiegnu* [NumP *tri* [NP [N' [#] [AP *nnivuri*.]]]]
 thereof= have.PRES.1SG three black.MPL
 “I’ve got three black ones.”
- (10) a. [DP [Spec *Ogni*] [D' D° ___ [NP [Spec *ppoveru*] *gattu*]]]
 every poor cat
 “Every poor cat”
- b. [TP *Un nni* [AspTermP [Spec *cchiù*] [AspConP...[AspPerfP[Spec
 not us= anymore
 ssempe] *chiama*.]]]]
 always call.PRS.3SG
 “He no longer always call us.”

By the same token, we predict that RF should not obtain whenever Word₁ and Word₂ surface in distinct phasal domains, a prediction borne out by contrasts like (11a–b) and (12a–b).

- (11) a. [TopP [Spec *Accussi*] [Top'... [TP *pparava* *ccu mia*.]]]
 thus speak.PST.3SG with me
 “He used to speak to me like that.” (Spec-Head)

- b. [_{ForceP} [_{Spec} *Accussì*] [_{Force'}... [_{TP} *parrava* *ccu mia*.]]]
 thus speak.PST.3SG with me
 “Therefore he used to speak to me.”
- (12) a. [_{ForceP} [_{Force'} *Fa-* [_{FinP} [_{Fin'} [_{Fin°} [_{Cl} *llu*] [_{Fin} *fa*<sub>[v-VP *lu fa!*]]]]]]]]]
 do.IMP.2SG =it.ACC.MSG
 “Do it!” (Head-Head)</sub>
- b. [_{ForceP} [_{Force'} *Fa* [_{FinP} [_{Fin'} *fa* [_{TP} ... *tuttu* [_{v-VP} *fa!*]]]]]]]
 do.IMP.2SG all
 “Do everything!”

In (11a) *accussì* receives a topical reading, namely “like that”, and *qua* modal adverb lexicalizes Spec,Top contained within the same (lower) CP phase as the finite verb *pparava* raised to the T-domain and bearing RF. In (11b), by contrast, the verb no longer bears RF and *accussì* is now a discourse connector interpreted as the resultative adverb “therefore” external to the clause. Under this parenthetical use (cf. Cinque 1999: § 1.6), we assume that the adverb is merged in a high C-related position, in (11b) labelled as Spec,Force for expository simplicity (but see Corr 2017 for detailed discussion of the syntax of the utterance beyond the canonical CP-layer), and, in particular, in a higher phasal CP domain than that in which the finite verb surfaces. As a consequence, RF fails to obtain since adverb and verb are contained within distinct phases and are sent to Spell-Out in distinct cycles of the derivation to PF (cf. also (2a) vs (3a) above).

A similar contrast is seen in (12a–b). As already noted in the discussion of (7b) above, in (12a) the imperatival verb excorporates from *Fin°* to reach a higher C-related (e.g., *Force°*) head, a movement we assume to also characterize (12b). However, only in the former example does the imperatival verb license RF (for identical facts in the dialect of Campobasso, see D’Ovidio 1874: 179–180), a difference that can be straightforwardly derived from the Phase Impenetrability Condition (PIC; Chomsky 2001, 2008). In (12a) the imperatival verb and the enclitic *lu* surface in distinct phases instantiated by *ForceP* and *FinP*, respectively, but the verb continues to license RF on its pronominal complement since the latter occurs in the left edge (*viz.* head) of *FinP* and hence remains accessible to the verb with which it is sent to PF in the same higher cycle. In (12b), by contrast, the verbal complement *tuttu* occurs in the lower portion of the pre-*v-VP* space contained within the complement of the lower *FinP* phase, hence inaccessible to the potential RF effects of the verb in the higher *ForceP* phase.

Note, finally, that in (12b) the intermediate position (*viz.* copy) of *fa* in *Fin°* is not able to trigger RF on *tuttu*, despite both items occurring within the same phase at this point in the derivation. This is a general property of unpronounced copies which invariably fail to license RF since they do not survive to PF, but,

rather, represent simple bundles of formal features which, while playing a role in the conceptual-intentional system, do not play any role in the sensory-motor system to which they are invisible once all their phonological features have been stripped away.

3. Pragmatico-semantic effects of RF

Having observed how the distribution of RF falls under the locality conditions imposed by phase theory and, in particular, the PIC, in the following final sections I examine two cases where the distribution of RF is shown to interact with clause structure to license and mark at the interfaces different pragmatic and semantic interpretations.

3.1 Pragmatic mapping

Above in § 2.1 we noted how all 3SG verbs systematically qualify as RF-triggers on account of the presence in their underlying phonological representation of a final consonantal slot (< -T). By way of illustration, consider the examples in (13).

- (13) a. *Iddu chiama nnu taxi.*
 he call.PRS.3SG a taxi
 ‘‘He calls a taxi.’’
- b. *Iddu si chiama Ciccii.*
 he self= call.PRS.3SG Ciccii
 ‘‘He’s called Ciccii.’’
- c. *Iddu un chiama mmai / ssempe / ttantu / ddumani.*
 he not call.PRS.3SG never always so.much tomorrow
 ‘‘He won’t call ever / always / very much / tomorrow.’’

In (13a–b) the verb is followed by nominal and predicative complements, respectively, a surface Head-Complement configuration which licenses the observed RF on the complement in both cases. In (13c), by contrast, the relationship between the finite verb and the following constituent is much looser, inasmuch as the latter represents in all cases an adjunct but, nonetheless, continues to bear RF, ultimately indicating that verb and adjunct are sent to PF in the same cycle. In other cases, however, RF is not invariably realized on the constituent immediately following a 3SG verb, witness the near minimal pair in (14a–b):

- (14) a. *Quannu vena ra cameriera, diciaccillu*
 when come.PRS.3SG the cleaner tell.IMP2SG=DAT.3SG=ACC.3MSG
ca ...
 that
 “When the cleaner arrives, tell her to ...”
- b. *Quannu vena *(r)a primavera, mi sientu ggìa*
 when come.PRS.3SG the spring me=feel.PRS.1SG already
cchiù mmiegliu.
 more better
 “When(ever) spring comes, I already start to feel much better.”

Surprisingly, the definite article in (14b), but not in (14a), cannot occur in its fortis variant with the initial vibrant,¹⁰ but assumes its simple lenis vocalic realization. However, the pragmatic reading of the two immediately postverbal subject DPs is distinct in both cases. In the first example the definite DP licenses a fully referential reading, identifying a specific and known referent salient in the discourse or the extra-linguistic context which we can characterize as topical. In the second example, by contrast, the DP, although definite, is not referential but, rather, receives a generic interpretation, hence the unbounded reading of *quannu* “whenever” in (14b) in contrast to its bounded interpretation “when” in (14a). Indeed, this referential difference is supported by examples like (15) where we now see that *quannu* has its bounded interpretation and the postverbal definite DP subject, now marked by RF, is concomitantly fully referential, and hence topical.

- (15) *Quannu vena ra primavera, m’ affittu na casa a*
 when come.PRS.3SG the spring me= rent.PRS.1SG a house at
ru mare.
 the sea
 “When the (= this) spring comes, I’ll rent a house by the sea.”

Further investigation of postverbal subjects shows, however, that it is not only topical subjects like (14a) and (15) which are marked by RF, but also focal ones. Consider the examples in (16a–b).

- (16) a. *Cchi succede? – VENA *(R)U POSTINU.*
 what happen.PRS.3SG come.PRS.3SG the postman
 “What’s happening?” – “The postman’s coming.”
- b. *Chine vena? – Vena *(R)U POSTINU.*
 who come.PRS.3SG come.PRS.3SG the postman
 “Who’s coming?” – “(It’s) the postman (who)’s coming.”

10. The vibrant represents the outcome of a weakening process of the original Latin long lateral, viz, -[ll]- > -[dd]- > -[qd]- (> -[jj]-) > -[r/r]- (cf. Ledgeway 2016: 254).

Here we see that, whether the postverbal subject occurs in wide (16a) or narrow (16b) focus, it invariably surfaces marked by RF. What Examples (16a–b) thus share with (14a) and (15b), but not with (14b), is that the definite DP subject is fully referential, irrespective of its topical or focal interpretation. Indeed, that the fully referential nature of the definite DP is at stake is also shown by examples such as (17a–b).

- (17) a. *Cchi ffa ru postinu?*
 what do.PRS.3SG the postman
 “What’s the postman doing?”
 b. *Picchi chiangia ru postinu?*
 why cry.PRS.3SG the postman
 “Why is the postman crying?”

Hitherto all examples have involved an immediately postverbal definite DP subject with an unaccusative verb, whereas the examples in (17a–b) involve transitive and unergative predicates, respectively. Structurally, the subjects of the former are known to represent internal arguments, whereas subjects of transitives and unergatives represent external arguments at all levels of representation. Despite these deep structural differences, postverbal subjects of transitives and unergatives behave just like postverbal unaccusative subjects in all relevant respects, highlighting how the distribution of RF in examples such as (17a–b) is licensed by the referential interpretation of the subject. This is further demonstrated by the corresponding examples in (18a–b) in which the postverbal subject now appears without RF.

- (18) a. *Cchi ffa u postinu?*
 what do.PRS.3SG the postman
 b. *??Picchi chiangia u postinu?*
 why cry.PRS.3SG the postman

An acceptable answer to the question in (18a) might be, for example, *Porta ri lettere* ‘He delivers letters’, inasmuch as the interpretation of the postverbal definite DP subject without RF in (18a) is not referential, but refers to the ‘role of the postman’, hence the generic reading *What do postmen do/does a postman do?* By contrast, the RF-marked *ru postinu* in (17b) is fully referential and refers to a specific, known individual who is a postman, giving rise to the observed presuppositional reading *What is the postman doing/up to?* Consequently, the latter can be uttered when enquiring about your local postman, an individual you see on a regular basis who is known to you, whereas the former can only be employed to enquire about the generic duties of postmen in general. Now, although (18a) is perfectly grammatical, example (18b) is generally judged to be unacceptable. However, the less than perfect status of the latter is not due to structural reasons, but follows, rather, from

the difficulty of finding a pragmatically acceptable context in which to utter it. In particular, when faced with such an example speakers say that, given our contextual use of semantic and real world knowledge about postmen, the question in (18b) proves pragmatically odd since there is no expectation that an intrinsic part of the postman role involves crying. In short, an example such as (18b) only licenses a generic, non-referential reading of *u postinu*, leading to the very unlikely question *Why do postmen cry?* This contrasts with the RF-marked referential use of *ru postinu* in (17b) where the speaker enquires about the individual act of crying of a particular postman.

Finally, consider the examples in (19)–(20):

- (19) a. *U muratore minava ??(r)u cane.*
 the builder beat.PST.3SG the.MSG dog.M
- b. *U muratore u minava *(r)u cane.*
 the builder him.ACC= beat.PST.3SG the.MSG dog.M
 “The builder was hitting the dog.”
- (20) a. *Maria si (a) mangia ra carne.*
 Maria self= it.FSG= eat.PRS.3SG the.FSG meat.F
 “Maria is eating the meat.”
- b. *Maria (si) *(a) mangia a carne (u marti).*
 Maria self= it.FSG= eat.PRS.3SG the.FSG meat.F the Tuesday
 “Maria eats meat (on Tuesdays).”

Example (19a) in which the postverbal definite DP object bears RF can *a priori* be interpreted both as a (wide/narrow) focus (in response to the underlying questions *What was the builder doing?/What was the builder hitting?*, respectively) or as a topic (in response to the question *What was the builder doing to the dog?*), although in the latter case the variant (19b) is generally preferred in which the topicalized DP object is clitic-doubled. The variant without RF in (19a), by contrast, proves once again pragmatically odd, since the object is now only capable of receiving a generic, non-referential interpretation and therefore not amenable to clitic-doubling (cf. (19b)). Consequently, the version of (19a) without RF on the direct object can only be interpreted to mean something like *The builder was a dog-hitter* where, in a particular world, part of the defining characteristics or duty of builders would be their propensity to hit dogs. Turning finally to the examples in (20), we note that in (20a) the definite DP object which bears RF receives a fully referential interpretation, hence the possibility once again of clitic-doubling, licensing a punctual eventive interpretation of the sentence in which the speaker identifies a particular act of eating of a specific piece of meat. In (20b), by contrast, the definite DP object appears without RF and now receives a non-referential interpretation (hence the ungrammaticality of clitic-doubling, and the preference for

the absence of the reflexive clitic *si*), giving rise to the generic/habitual reading of the act of meat-eating.

To sum up, the distribution of RF reviewed above clearly shows that the dialect of Cosenza formally distinguishes between postverbal definite referential DPs, be they topical or focal, and their non-referential variants. Given our previous conclusion that for RF to take place *Word*₁ and *Word*₂ must surface in the same phasal domain, we now have a principled explanation for the facts considered above. In particular, adopting Belletti's (2004, 2005) seminal idea that the *v*-VP edge makes available a lower left periphery with dedicated Topic and Focus positions, we assume a direct mapping between syntax and pragmatico-semantic interpretation such that all referential constituents, when not raised to the higher left periphery, target a Topic or Focus position within the lower left periphery, whereas all non-referential constituents remain in situ within the VP (cf. Diesing's 1992 Mapping Hypothesis).¹¹ Consequently, we associate a minimal pair such as *Quannu vena (r)u postinu?* with the structural representations in (21), where the presence of RF on the postverbal definite DP in (21a) signals a referential reading of the subject raised to Spec,Top, namely *When is the (= our) postman coming?*, whereas its absence in (21b) correlates with a non-referential interpretation of the definite DP in situ, namely *What time is the postal delivery?*

- (21) a. *Quannu vena* [_{TopP} [_{Spec} *ru postinu*] [_{vP} *vena* [_{vP} *vena u postinu?*]]]
 when come.PRS.3SG the postman
- b. *Quannu vena* [_{TopP} _____ [_{vP} *vena* [_{vP} *vena u postinu?*]]]
 when come.PRS.3SG the postman

Following Ledgeway & Lombardi (2005), I take the finite verb in Cosentino to target a low functional head situated above the *v*-VP complex. It therefore follows that RF is licensed with referential postverbal subjects such as (21a) where the finite verb (*viz.* *Word*₁) and the immediately postverbal constituent (*viz.* *Word*₂) are transferred to PF in the same higher cycle, since the postverbal subject surfaces in the left edge of the lower *v*P phase from where, in accordance with the PIC, it remains accessible to phonosyntactic processes of the higher CP phase. In (21b), by contrast, the postverbal subject from its in situ position remains inaccessible to the potential RF effects of the 3SG finite verb, since it is contained within the *v*P phase from where it is sent to PF in the lower cycle before the spell-out of the RF trigger in the higher phasal cycle.

11. There are obvious parallels here with those analyses of Romance DOM which assume a direct mapping between interpretation and narrow syntax, with raising of presuppositional DPs to the *v*-VP edge (cf. Torrego 1998; Ledgeway 2000: Chapter 2).

3.2 Coreference

I finally turn to consider the interaction of the distribution of RF and coreference. I begin by considering the examples in (22).

- (22) a. *Iddu_i sa / dicia / crida cca pro_{i/??j} puzza.*
 he know.PRS.3SG say.PRS.3SG believe.PRS.3SG that smell.PRS.3SG
 “He_i knows / says / believes that he_{i/??j} smells.”
- b. *Iddu_i sa / dicia crida ca pro_{j/*i} puzza.*
 he know.PRS.3SG say.PRS.3SG believe.PRS.3SG that smell.PRS.3SG
 “He_i knows / says / believes that he_{j/*i} smells.”

Although in both examples the main verb selects for a propositional CP complement introduced by the complementizer *ca* “that”, only in the first example does the latter show RF. This difference is not without its consequences but, rather, serves to mark an important structural distinction between these otherwise apparently identical examples. Specifically, in (22a) where RF obtains, the most natural and universally accepted reading is one in which the null embedded subject is obligatorily interpreted as coreferential with the matrix subject, although for some speakers the disjoint reading is reported to be marginally possible.¹² However, all speakers agree that the usual way of marking the disjoint reading is as in (22b), where RF on the finite complementizer fails to obtain and the coreferential reading is universally rejected by all speakers.

This result is further evidenced by the examples in (23) where once again the coreferential reading can only be licensed when the complementizer bears RF as in (23a), such that a full DP subject with disjoint reference (e.g., *Pinu*) can only be entirely felicitously licensed for all speakers in (23b) where the complementizer does not bear RF.¹³

12. Diagenational factors also seem to be at play here, inasmuch as RF under the disjoint reading is entirely rejected by older speakers. Consequently, the marginal status of RF under the disjoint reading seems to represent an incipient innovation in the grammars of younger speakers who (over)generalize RF to all occurrences of the finite complementizer immediately following a 3SG finite verb.

13. These same results are confirmed by examples such as (i) involving split antecedence (viz. non-exhaustive control) where RF is once again heavily dispreferred on account of the partially disjoint interpretation of the matrix and embedded subjects.

- (i) *Cicciu_i (mi_j) prummintia (??c)ca pro_{i+j} putiamu viaggia nzieme.*
 Cicciu me= promise.PST.3SG that be.able.PST.1PL travel.INF together
 “Cicciu_i was promising (me_j) that we_{i+j} could travel together.”

- (23) a. *Franco_i dicia cca pro_{i/??j} / ?Pinu tena raggiune.*
 Franco say.PRS.3SG that Pinu have.PRS.3SG reason
 “Franco_i says that he_{i/??j} / ?Pinu is right.”
- b. *Franco_i dicia ca pro_{j/*i} / Pinu tena raggiune.*
 Franco say.PRS.3SG that Pinu have.PRS.3SG reason
 “Franco_i says that he_{j/*i} / Pinu is right.”

In light of this evidence, we should expect that irrealis subjunctive complement clauses, also introduced by the finite complementizer *ca* in the modern dialect (cf. Rohlfs 1983: 152; Ledgeway 2009b), should pattern with propositional complement clauses involving disjoint reference since the former, on account of the obviation effect, typically involve disjoint reference between matrix and embedded subjects.¹⁴ Indeed, this prediction is borne out by examples such as (24a–b).

- (24) a. *Rosina_i vulia ((?c)ca pro_j pagassaru subbitu.*
 Rosina want.PST.3SG that pay.PST.SBJV.3PL at.once
 “Rosina_i wanted them_j to pay at once.”
- b. *Iddu_i s’ aspetta ((?c)ca pro_j ni manna na foto.*
 he self= waits that us= send.PRS.3SG a photo
 “He_i expects him_j/her_j to send us a photo.”

From a theoretical perspective these facts lend themselves to and, at the same time, provide further support for, a number of current ideas about the nature of phases. In particular, we have seen that where subject coreference between matrix and embedded subjects obtains the head of the embedded clause, viz. *ca*, bears RF, whereas RF is not licensed whenever matrix and embedded subjects are (potentially) disjoint. Intuitively, what these data seem to suggest is that there is greater interlacing and structural integration between matrix and complement clauses under the

14. Of course, in subjunctive complement clauses the embedded subject can be exhaustively coreferential with a matrix object (cf. ia), but partial coreference is also allowed in the same contexts (cf. ib), highlighting how it is the absence of obligatory exhaustive coreference in such contexts which precludes RF on the finite complementizer:

- (i) a. *A Rosina_j mamma cci_j raccumannava *(c)ca pro_j si*
 to Rosina mum 3.DAT recommend.PST.3SG that self=
ricugliassa a menzanotte.
 bring.SBJV.3SG at midnight
 “Rosina_i, mum warned her_i to return home by midnight.”
- b. *A Rosina_i mamma cci_i raccumannava *(c)ca pro_{i+j} si*
 to Rosina mum 3.DAT recommend.PST.3SG that self=
ricugliassaru a menzanotte.
 bring.SBJV.3PL at midnight
 “Rosina_i, mum warned her_i that they_{i+j} should return home by midnight.”

coreferential reading (Lehmann 1988), paralleled, in turn, by a commensurate degree of semantico-pragmatic integration between the two events (Givón 1990: 826), than under the disjoint reading where the complement clause appears to enjoy greater autonomy and display less integration with the matrix clause.

Drawing on parallels with recent work by Sheehan (2014, 2018a; b) on exhaustive and partial control in Romance and beyond, we can distinguish between complement clauses such as (22a) and (23a) where the embedded subject is obligatorily null and exhaustively coreferential with a local matrix subject and cannot be substituted by an independent pronoun or lexical DP, and complement clauses such as (22b), (23b) and (24a–b), as well as examples (i) and (i.a–b) in fns 13 and 14, where the embedded subject position does not enter into any such relation of obligatory exhaustive coreference and may freely refer, partially or exhaustively, to a (non-local) matrix argument or be realised by a(n overt) (pro)nominal with independent reference.¹⁵ In the former case the embedded T, although inflected for agreement, arguably fails to license nominative, witness the obligatory coreferential subject reading derived via A-movement,¹⁶ whereas in the latter case embedded T licenses nominative Case variously yielding the observed (partially) coreferential and disjoint readings. Following Chomsky (2001, 2012), (strong) phases are those syntactic domains in which all features have been valued, including phi- and Case features of the subject (cf. also Richards 2007; Gallego 2010; Uriagereka 2011), such that we take cased NOC and uncased OC complement clauses to be (strong) phases and (weak) non-phases, respectively.

Returning then to contrasts like (22a–b), it is tempting to interpret the observed distribution of RF on the complementizer as an effect of phase theory: on the assumption that as referential arguments both CP complements in (22) must raise to a position within the lower left periphery (cf. (21a), then the RF on the complementizer in (22a) follows straightforwardly, since the complement clause is not a phase such that the complementizer lexicalising the head of ForceP (cf. Ledgeway 2009b) is accessible and sufficiently local to the RF trigger instantiated by the matrix verb in the lower pre-*v*P field. However, by the same token, it is also incorrectly predicted that the complementizer in (22b) will display RF since, although the complement clause is phasal, the complementizer occurs in its edge (viz. in Force^o) and hence, in accordance with the PIC, continues to be accessible

15. In what follows, we will informally refer to the former as obligatory control (OC) complement clauses, and the latter as non-obligatory control (NOC) complement clauses.

16. For arguments that obligatory exhaustive coreference in apparently finite clauses should be dealt within as a case of A-movement along the lines of OC control into infinitival complements, see the extensive discussion in Ledgeway (2000: Chapter 3).

and sufficiently local to the relevant RF trigger.¹⁷ One could of course try to save the underlying intuition that the phasal nature of NOC complements in some sense blocks RF by arguing that *ca* lexicalizes a lower C-head such as *Fin*^o (on irrealis clauses, see discussion in fn. 17), as represented in (25a). Empirically, however, this analysis is not tenable, as shown by examples such as (25b) where the embedded left periphery hosts a topicalized constituent but the complementizer *ca*, although preceding the fronted topic in Spec,TopP, hence arguably lexicalizing Force^o, still fails to display RF.

- (25) a. *Cicciu*_i *dicia* [_{ForceP}... [_{FinP} (**c*)*ca* [_{TP} *pro*_j / *Pinu a sbagliatu i cuonti*]]].
 Cicciu says that Pinu have.PRS.3SG
sbagliatu i cuonti]]].
 mistake.PTCP the accounts
 “Cicciu_i says that he_j/Pinu got the sums wrong.”
- b. *Cicciu*_i *dicia* [_{ForceP} (**c*)*ca* [_{TopP} [_{Spec} *i cuonti*] [_{TP} *pro*_j / *Pinu l'a sbagliati*]]].
 Cicciu says that the accounts Pinu
l'a sbagliati]]].
 them=have.PRS.3SG mistake.PTCP.3MPL
 “Cicciu says that, as for the sums, he/Pinu got them wrong.”

Although we cannot appeal therefore to phasehood as a potential locality intervention effect inhibiting the application of RF, I argue that the relevant differences in the distribution of RF can still be derived from phase theory. In particular, I return to a traditional idea about the licensing of clausal arguments, namely Stowell’s (1981) Case Resistance Principle (CRP) which states that Case may not be assigned to a category bearing a Case-assigning feature. Building on this intuition, we have already seen that NOC complements are cased strong phasal domains in that T assigns nominative to the embedded subject position, thereby identifying such clauses as arguments which resist Case-licensing under the CRP. By contrast, OC complements are uncased weak phasal domains which require the embedded subject to raise to the matrix clause since embedded T does not bear a nominative Case-assigning feature, thereby singling out such clauses as subject to Case-licensing in line with the CRP. In this way, the distribution of RF can be derived quite simply from the CRP and, ultimately, phase theory: while NOC

17. Furthermore, this analysis presents the further complication that it incorrectly predicts a difference between NOC realis and irrealis complement clauses. In contrast to propositional complement clauses where realis *ca* lexicalizes Force^o, Ledgeway (2009b) shows that in modern Cosentino *ca* lexicalizes the lower *Fin*^o head in irrealis complement clauses, hence situated outside of the highest C-related phasal edge. This would therefore incorrectly lead us to expect in line with general assumptions about the PIC a distinction between NOC realis and irrealis clauses with RF surfacing on *ca* in the former but not the latter case.

complement clauses are characterized by a greater degree of autonomy and finiteness (viz. T[+Nom]) licensing, in Stowell's terms, an intrinsic structural Case feature in C° which places such clauses outside of the usual Case-theoretic licensing conditions (Bošković 1995) and forcing them, in accordance with Stowell's original proposal, to surface in extraposed positions, OC complement clauses, on account of their reduced autonomy and finiteness (viz. T[-Nom]), have to be Case-licensed.¹⁸ My proposal then is quite simply that, unlike NOC complement clauses which are obligatorily extraposed under the CRP (cf. (26a), OC complement clauses (cf. (26b) are Case-licensed within the *v*-VP complex and raise like other referential arguments (cf. (21a) to the lower left periphery.

- (26) a. *Luca_i sa* [_{v-VP} *sa* [_{ec_k}] [(**c*)*ca pro_j sgarra*]_k].
 Luca know.PRS.3SG that err.PRS.3SG
 "Luca_i knows that he_j is wrong."
- b. *Luca_i sa* [_{FocP} [_{Spec} *cca Luca_i sgarra*]
 Luca know.PRS.3SG that err.PRS.3SG
 [_{vP} *sa* [_{VP} *sa* [_{ca} *Luca_i sgarra*]]]]].
 "Luca_i knows that he_j is wrong."

The observed RF facts now follow without further stipulation: extraposition of the NOC complement clause in (26a) places it in a different phasal domain (viz. the *vP*) than the potential RF trigger, the finite verb *sa* contained within the higher CP phase, whereas in (26b) the OC complement clause surfaces in the left edge of *vP* and hence is sent to PF in the same phasal cycle as the verbal RF trigger in accordance with the PIC. We thus observe that the distribution of RF is not constrained by the presence of a potential phasal barrier interrupting the immediate structural contiguity between Word₁ and Word₂, but, rather, is a direct consequence of the different phasal status of the two clausal complement types. In accordance with the original generalizations underlying the CRP, this difference results in the extraposition of all phasal CPs placing them beyond the potential RF effects of a selecting 3SG verb, in contrast to their non-phasal counterparts which undergo raising to the *vP* left periphery as part of their licensing. Indeed, this conclusion finds further structural support in the observation that whereas extraposed phasal CPs can be readily doubled by a resumptive object clitic (cf. (27a), the same is not true of non-phasal CPs (cf. (27b).

18. Relevant here are the frequent claims in the literature (cf. Massam 1985; Plann 1986; Raposo 1987; Acquaviva 1989; Ledgeway 2000: Chapter 3, 6) that non-finite CPs in Romance require Case-licensing.

- (27) a. *Luca_i u_k sa [ec_k] [ca pro_j sgarra]_k.*
 Luca it= know.PRS.3SG that err.PRS.3SG
 “Luca_i knows that he_j is wrong.”
- b. *Luca_i *(u_k) sa [cca ~~Luca_i~~ sgarra]_k.*
 Luca it= know.PRS.3SG that err.PRS.3SG
 “Luca_i knows that he_j is wrong.”

In conclusion, the presence or absence of RF on the complementizer falls out from standard interpretations of locality computed in terms of the phasal domains which we have seen apply in all other potential contexts of RF.

4. Further implications and extensions

Although my investigation of the distribution of Cosentino RF has shown how the phonological and syntactic components may operate in unison, with phonological domains invariably aligning with syntactic domains to externalize at PF various syntactic constraints and distinctions, the distribution of RF in many other dialects reveals that such an isomorphic mapping between syntax and PF represents just one of several parametric possibilities at the interfaces. Indeed, the distribution of RF in the dialects of southern Italy, which display considerable subtle but structured microvariation in the licensing and surface effects of RF, offers many valuable opportunities to deepen our understanding of the phonological correlates of syntactic representations. In the case of Cosentino we have seen how an isomorphic approach to the mapping of syntactic Spell-Out domains on phonological domains makes the right predictions about the distribution of RF in this variety. Similarly, on the basis of a small sample of southern dialects, Ledgeway (2018: § 3) sketches how the structurally-determined alternations in the distribution of RF following the singular persons of the perfective auxiliary and copular uses of BE (viz. 1/2/3SG SUM/SIS/EST > *so/si/è*),¹⁹ widespread in southern dialects (cf. Manzini & Savoia 2005, II–III: Chapters 5–6; Torcolacci 2014a; b; Ledgeway 2019), can, in principle, also be handled in terms of a phase-theoretic approach in which syntactic and phonological domains are taken to align, albeit in conjunction with other assumptions such as parametric variation in verb movement.

However, an examination of a larger selection of southern dialects like those presented in Table 1 gleaned from Manzini & Savoia (2005) immediately reveals that such an isomorphic approach to syntactic and phonological domains is not tenable in all cases.

19. The singular persons of BE are variously realized in accordance with local phonetic variation as [so, sɔ, su] (1SG), [si, fɪ, sje] (2SG) and [e, e, i]. In many dialects, the 1SG forms are also often syncretic with the 3PL (cf. the dialects of Capracotta and Vastogirardi discussed below).

Table 1. Distribution of RF following singular persons of BE*

	Transitive			Reflexives			Unaccusative			Copula		
	1	2	3	1	2	3	1	2	3	1	2	3
Poggio Imperiale	+	+	+	+	+	+	+	+	+		+	+
Guardiaregia		+			+	+		+	+		+	+
S.Giorgio del Sannio			+			+			+	+	+	+
Bitetto			+			+		-	+	+	-	+
Frigento			+						+		-	+
Capracotta		-	-/+	-	-	-/+	-	-	-/+	-	-	-/+
Putignano				+	+	+	+	+	+	+	?	+
S.Lorenzo del Vallo						+		-	+	?	?	+
Giurdignano							+	+	+	+	?	+
Maglie							+			+		
S.Vittore		-			-				-	+	-	+
Molfetta	-	-		-	-		-	-		-	-	+
Amandola	+	+		?	?	?	+	+	+	+	-	+
Sonnino	+	-		+	-		+	+		?	?	+
Viticuso		+			-			+		+	+	+
Padula			+			-	+	-	+	?	?	?
Vastogirardi		-	-/+		-	-/-		-	-/+		-	-/-
Roccasicura	-	-	-	-	-	-	-	-	+	-	-	+
Pontecorvo	-	+		-	+		-	+		+	+	+
Secinaro	-	+		-	?	+	-	+	+	+	+	+
Castelvecchio Subequo	-	+		?	?	?	+	+	+	+	+	-
Ruvo di Puglia	-	-	-	+	-	+	+	-	+	+	-	+
Gravina di Puglia	-		+	-		-	+		+	+	-	-
Popoli	+	+		-	?	+	-	-	+	+	-	+
Agnone	-	-		-	-	+	-	-	-	+	-	
S.Benedetto del Tronto	-	-		?	?	?	-	-		+	+	+
Tuffillo	-	-		?	?	?	+	+	+	?	?	?
Pàstena-Castelpetroso	-	-	-	-	+	+	+	+	+	+	-	+

* Table 1 distinguishes between auxiliary and copular uses of BE in which it is respectively followed by a transitive/unergative, reflexive and unaccusative participle and by an adjective. ± indicate the presence/absence of RF, ? the absence of relevant information in Manzini & Savoia (2005), and an absence of any symbol indicates that the relevant potential RF trigger is not licensed in that particular person/context.

Although the data reported in Table 1 are far too numerous and complex for us to be able to treat them exhaustively here, we can nonetheless draw out some provisional yet significant patterns. The first 12 varieties contained in the first section of Table 1 (henceforth Group 1) are those in which a given BE RF trigger, whenever available in one of the four given contexts, systematically licenses RF on the following active participle or, in its copulative uses, on the following predicative complement. Thus, although the individual varieties in Group 1 may vary enormously in terms

of the distribution of the relevant singular forms of BE including, for instance, the selection of all three singular forms of BE [+RF] in conjunction with all active participles in the dialect of Poggio Imperiale but limited to the participles of reflexive and unaccusative predicates in the dialect of Putignano, or the systematic RF effect of the 2SG form in all four contexts in the dialect of Guardiargia in contrast to the systematic absence of any such RF effect with the 2SG form in the dialects of Capracotta, San Vittore and Molfetta,²⁰ what binds them all together is the observation that, if marked as an RF trigger in a given variety, the relevant forms of BE invariably license RF whenever selected.

By contrast, the 8 varieties contained in the second section of Table 1 (Group 2) include a single contrast in one of the three persons, whereby the potential RF trigger fails to license RF in one or more of the four contexts whilst triggering it in one or more of the remaining contexts. For example, in contrast to the 2/3SG forms which consistently license RF in the dialect of Secinaro, the 1SG systematically fails to license RF on a following active participle, but does trigger it on a following adjectival complement.

The 7 dialects in the third section of Table 1 (Group 3) introduce a contrast in two of the persons. For instance, in the dialect of Castelvecchio Subequo there is both a contrast in the 1SG which triggers RF with unaccusative participles and adjectival complements but not with transitive participles, and in the 3SG which licenses RF in conjunction with a following unaccusative participle, but not with an adjectival complement, whereas the 2SG consistently licenses RF in all contexts in which it occurs.

The fourth and final section of Table 1 contains just 1 dialect (Group 4) where all three grammatical persons each engender a contrast. Consequently, in the dialect of Pàstena-Castelpetroso the 1SG triggers RF before unaccusative participles and adjectival complements but not before transitive and reflexive participles, while the 2SG licenses RF only in conjunction with reflexive and unaccusative participles, and the 3SG licenses it everywhere except before transitive participles.

While further detailed investigation of the individual microparametric properties of each of the varieties contained in each of the four macrogroups in Table 1 is necessary, it is abundantly clear that a phase-theoretic approach which assumes a strict alignment in the mapping between syntactic and phonological domains, though able to adequately account for the consistent behaviour of RF triggers in varieties such as Cosentino, fails to explain the increasingly less consistent and

20. I have exceptionally included the dialect of Capracotta in Group 1 since, although all singular forms of BE consistently fail to license RF (cf. discussion of the dialects of Monteroduni and Gallo below), the 3PL form *so*, despite being homophonous with the 1SG form, invariably licenses RF whenever selected. Compare the similar case of Vastogirardi in Group 2.

unpredictable behaviour of potential RF triggers in Groups 1, 2, 3 and 4. Take for example the dialect of Pontecorvo where we witness a voice split, in that the 1SG form of BE never triggers RF on a following active participle,²¹ but consistently licenses RF in conjunction with a stative adjectival complement (28a). On the plausible assumption that active *v*Ps are phasal but non-active *v*Ps are not, the observed RF contrast follows straightforwardly since only the complement of BE containing an active participle is sent to PF in an earlier phasal cycle than that containing BE. By contrast, in the 2SG we find almost the opposite distribution of RF with the relevant form of BE licensing RF in conjunction with active participial complements, which are expected to block RF on account of their strong phasal status, as well as with stative adjectival complements (28b).

Pontecorvo, southern Lazio (Manzini & Savoia 2005, II: 701–702)

- (28) a. (mə) su par'lacə / la'vacə / və'nucə / kkun'təŋcə.
 me= be.PRS.1SG spoken washed come happy
 "I have spoken / washed myself / come // I am happy."
- b. (tə) si ppar'lacə / lla'vacə / vvə'nucə / kkun'təŋcə.
 you.2SG= be.PRS.2SG spoken washed come happy
 "You have spoken / washed myself / come // You are happy."

All things being equal, the observed difference in the distribution of RF following the 1SG and 2SG person forms of BE in the dialect of Pontecorvo is therefore difficult to reconcile within the traditional isomorphic approach to phases, since the expected syntactic restriction is reflected in the phonological output only in conjunction with the 1SG, but not in conjunction with the 2SG.

Conflicting data like these, of which there are many more examples in Table 1, highlight that it is not always possible nor necessary to postulate a direct mapping of syntactic Spell-Out domains onto phonological domains, inasmuch as the two may operate in isolation. Indeed, data like these lend strong empirical support to D'Alessandro & Scheer's (2015) non-isomorphic approach to Phase Theory in terms of a modular interpretation of the PIC. In particular, they propose that, in accordance with parametric variation, individual phase heads may variously be associated with a PIC either in the syntax or the phonology, in both or in neither. This modular approach to the PIC therefore gives rise to the four logical combinations in Table 2 based on the typology presented in D'Alessandro & Scheer (2015: 602).

21. Note that in this variety reflexive predicates align with transitives, but in other dialects, as we shall see below, they may align with unaccusatives. For further discussion of the variable behaviour of reflexive predicates, see the discussion in Ledgeway (2019: § 4.3).

Table 2. Modular approach to PIC (cf. D'Alessandro & Scheer 2015: 602)

PIC in syntax	PIC at PF	Application of RF	Example
+	+	No	Cosentino
+	–	Yes	S.Giorgio del Sannio, Bitteto
–	+	No	S.Benedetto del Tronto
–	–	Yes	Cosentino

The first possibility represented in the top row of Table 2 partially characterizes varieties such as Cosentino where we have seen that, whenever a given phase head is endowed with a PIC in the syntax, it is also systematically associated with a corresponding PIC at PF, thereby blocking RF since there is an isomorphic alignment between syntactic and phonological domains in all cases. By the same token, the fourth option in Table 1 also correctly identifies varieties like Cosentino, in that an absence of a PIC in the syntax equally presupposes an absence of a corresponding PIC at PF, thereby licensing RF in the relevant cases. As for the second option in Table 2, this accurately represents what we find in varieties where, despite a potential phasal boundary in the syntax, PF invariably treats BE and its complement as constituting a single phonological domain licensing RF in all cases. For example, in the dialects of S.Giorgio del Sannio and Bitetto the 3SG RF trigger indiscriminately licenses RF on all participial and adjectival complements irrespective of an active-stative phasal opposition in the syntax which distinguishes between transitives/unergatives on the one hand (considered to be strong phases) and unaccusatives and copular predicative complements on the other (considered to be weak phases).

The opposite behaviour is exemplified by the third option in Table 2 where a given phase head, even though not associated with a PIC in the syntax, is nonetheless endowed with a PIC at PF, such that the relevant forms of BE systematically fail to license RF despite the absence of a phasal boundary in syntax. This is the case, for example, of the dialect of S.Benedetto del Tronto where in the 1/2SG the relevant forms of BE license RF in conjunction with a stative adjectival complement, but not, surprisingly, in conjunction with an unaccusative participle despite its weak phasal status. As a result, unaccusatives pattern in this variety with transitives, which, on account of their strong phasal status, are also predicted to not display RF in conjunction with the 1/2SG forms of BE, ultimately revealing how the phonological contrast in the distribution of RF underlies an active vs non-active split in the syntax (cf. discussion of Ariellese above).

While the four options in Table 2 demonstrate how individual phasal heads and constructions may, in isolation, variously behave in relation to D'Alessandro & Scheer's (2015) modular interpretation of the PIC, an integrated formalization

of the available options is illustrated in Table 3 which naturally captures the full extent of cross-dialectal variation displayed in Table 1:

Table 3. Distribution of RF with singular persons of BE according to modular PIC

PIC in syntax	PIC at PF	Application of RF	Example
±	±	±RF	Cosentino, Putignano, S.Lorenzo del Vallo, Giurdignano, Maglie, S.Vittore, Molfetta, Ruvo di Puglia, Roccasicura, Tuffillo
±	–	+RF	Poggio Imperiale, Guardiaregia, S.Giorgio del Sannio, Bitetto, Frigento, Capracotta
±	+	–RF	Monteroduni, Gallo Matese
±	∓	±RF	Amandola, Sonnino, Vitucuso, Padula, Vastogirardi, Pontecorvo, Secinaro, Castelvecchio Subequo, Gravina di Puglia, Popoli, Agnone, S.Benedetto del Tronto, Pàstena-Castelpetroso

Unlike the featural representations in Table 2, those in Table 3 are intended to explicitly model the global parametric characterization of individual varieties, rather than individual phasal heads or constructions. Take again, for example, Cosentino where we know that the presence or absence of a PIC in the syntax is invariably matched by a corresponding presence or absence of a PIC at PF. In the representation in Table 2 there was no way of capturing this isomorphic distribution holistically, inasmuch as Cosentino displays two of the potentially unrelated available options which each contribute to half of the relevant empirical generalization. In Table 3, by contrast, the perfect alignment between syntax and phonology witnessed in Cosentino is naturally captured by the first option, where both the presence and the absence of a PIC in the syntax are matched by a corresponding presence and absence of a PIC at PF. This isomorphic characterization equally holds of a number of varieties listed in Table 1, including from Group 1 Putignano, S.Lorenzo del Vallo, Giurdignano, Maglie, S.Vittore and Molfetta where the relevant BE trigger only ever licenses RF with non-transitives that lack a PIC in the syntax and hence also at PF.²² Similarly, in the Group 3 dialect of Ruvo di Puglia RF in the 1/3SG is found in unaccusative environments (including in this case reflexives), but is inhibited before transitive participles that instantiate a phasal boundary in both the syntax and at PF, with an analogous pattern obtaining in the Group 2 and 3 dialects of Roccasicura and Tuffillo in the 3SG and 1/2SG, respectively.

22. Note, as anticipated in fn. 22, that reflexive participles in Putignano and S. Lorenzo del Vallo pattern with unaccusatives.

Moving on to the second option in Table 3, this non-isomorphic featural combination identifies those varieties where a given phase head, irrespective of whether it is associated with a PIC in the syntax, is never endowed with a PIC at PF, inasmuch as a BE trigger in these varieties invariably licenses RF despite the potential presence of a phasal boundary in syntax. This provides an accurate description of the first six varieties in Group 1 (namely, Poggio Imperiale, Guardiaregia, S.Giorgio del Sannio, Bitetto, Frigento and Capracotta) where the relevant forms of BE, whenever selected, indiscriminately license RF on all complement types alike, including transitive/unergative participles.

The third featural pairing in Table 3 arguably instantiates varieties such as Monteroduni (29a) and Gallo Matese (29b).

(29) Monteroduni, southwestern Molise (Manzini & Savoia 2005, II: 716–717)²³

- a. (tə) si / (s) ε rum'muitə / mə'nutə / la'vatə /
 you= be.PRS.2SG self=be.PRS.3SG slept come washed
 kun'diəndə.²³

happy

“You have / he has slept / come / washed your-/himself // You are / he is happy.”

Gallo Matese, southwestern Molise (Manzini & Savoia 2005, II: 717–718)²⁴

- b. (mə) sɔ / (tə) sə / (s) ε də'r'mu:tə/
 me= be.PRS.1SG you=be.PRS.2SG self= be.PRS.3SG slept
 mmə'nu:tə/ la'va:tə.²⁴
 come washed

“I have / you have / he has slept / come / washed my-/your-/himself.”

Although these two dialects display all the relevant forms of the potential RF BE triggers found in other surrounding dialects, they never license RF. Rather than instantiating an arbitrary lexical distinction according to which the singular forms of BE in these varieties would simply not be listed among the relevant RF triggers, these facts can be understood more naturally and elegantly in terms of D'Alessandro and Scheer's modular PIC. In particular, we take the complements

23. The 1SG form sɔ, a potential RF trigger, is uniquely found in copular constructions (alongside the variant bisyllabic form sɔŋgə) where it also fails to trigger RF, e.g., sɔ(ŋgə) kun'diəndə “I am happy”. Outside of copular constructions, only the 1SG bisyllabic form of BE (this time in free variation with the 1SG form of HAVE) is found.

24. Note that the data in Manzini & Savoia (2005) indicate that the participle “come” in the dialect of Gallo Matese displays some (inconsistent) lexicalized cases of initial consonantal lengthening (viz. [mm-]) such as in the example in (39b), which are not the result of RF (cf. also 1/2PL se:mə/ se:mə mmə'nu:tə “we have/ you have come”).

of auxiliary/copular BE in these varieties to be invariably endowed with a PIC at PF, irrespective of whether they are associated with a corresponding PIC in the syntax (as in the case of transitive participles) or not (as in the case of unaccusatives and copular complements). As in the previous case, then, the syntax once again fails to leave a footprint at PF which, endowed with a PIC, blindly blocks RF in all cases.

To conclude, we turn to the final featural pairing at the bottom of Table 3. Just as with the first option in Table 3, this option also makes reference to varieties where both syntax and phonology are variously associated with a PIC in different contexts, but in contrast to the former, the pairing is not (fully) isomorphic. By way of example, consider the Group 2 dialect of Amandola where the 1/2sg indiscriminately license RF with transitives and unaccusatives despite their respective strong and weak phasal characterizations in the syntax, but the 2sg alone unexpectedly fails to license RF in conjunction with a copular predicative complement. Such erratic distributions of RF underline how there is no necessary harmonic alignment of the PIC in syntax and at PF, with instantiations of individual phasal heads presenting unpredictable behaviours.²⁵ In short, in such cases PF-externalization is not directly fed by the narrow syntax, in that the disassociation between syntax and phonology produces a non-isomorphic mapping of the PIC at the interfaces.

Further examples from Table 1 which instantiate this non-isomorphic mapping between syntax and phonology include the remaining 6 dialects of Group 2 (namely, Sonnino, Viticuso, Padula, Vastogirardi, Pontecorvo and Secinaro), as well as the remaining 6 dialects of Groups 3 and 4 (namely, Castelvechio Subequo, Gravina di Puglia, Popoli, Agnone, S. Benedetto del Tronto and Pàstena-Castelpetroso). For instance, in the dialect of Sonnino the 1sg is not associated with a PIC at PF, with RF occurring in conjunction with transitive, reflexive and unaccusative participles alike, despite their differing phasal characterizations and associations with a PIC in the syntax. Yet, the 2sg which might a priori be expected to display an analogous distribution of RF unexpectedly licenses RF only in conjunction with an unaccusative participle, but not with transitive or reflexive participles, in line with an unpredictable and inconsistent active-stative split restricted to the 2sg. A near identical situation was also observed above in the dialect of Castelvechio Subequo, with the difference that the 2sg licenses RF with both transitive and unaccusative participles in contrast to the 1sg where RF is limited to unaccusative participles.

In both dialects, we are therefore forced to conclude that PF-externalization signals, at most, an inconsistent representation of an underlying syntactic footprint. An equally puzzling surface distribution is attested in the dialect of Gravina di

25. Indeed, the examples considered here precisely illustrate those options which Bonet et al. (2019: § 3.1) incorrectly characterize as unattested options overgenerated by an “excessively powerful” and “unrestricted” Modular PIC.

Puglia where the 1sg and 3sg give rise to two disparate and irreconcilable distributional patterns of RF: whereas the distribution of RF in conjunction with the 1sg appears to signal a classic active-stative split in the syntax with RF restricted to unaccusative and copular complements, RF in conjunction with the 3sg signals, by contrast, a syntactically unpredictable split between transitive and unaccusative participles marked by RF on the one hand, and reflexive and copular (participial) complements on the other which occur without RF.

5. Conclusion

The distribution of Cosentino RF has been shown to involve an isomorphic mapping of syntax and phonology at the interfaces, with phonological domains aligning with syntactic domains to externalize at PF syntactic information which, in turn, may spell out key semantico-pragmatic distinctions such as referentiality (cf. § 3.1) and coreference (cf. § 3.2). In particular, we have observed how the distribution of Cosentino RF is constrained by specific locality conditions which cannot be exhaustively computed in terms of the traditional core structural Spec-Head, Head-Head and Head-Comp configurations, but, rather, must be understood and modelled in terms of a unitary phase-theoretic approach. Not only does this approach offer a more adequate explanation of the structural restrictions on the application of Cosentino RF, but the distribution of the latter also provides new and interesting data to test the nature and computation of phasal domains, witness, for example, our conclusion that the postulation of a single CP phase is not sufficient to account for Cosentino examples such as (11)–(12) which force us to recognise distinct lower and higher phasal domains within the C-domain.

However, an examination of perfective auxiliary/copular BE RF patterns across a selection of southern dialects has highlighted how an isomorphic alignment of syntactic and phonological domains cannot invariably be assumed. Rather, the isomorphic mapping between syntax and PF at Spell-Out observed in Cosentino represents just one of four possible parametric pairings (cf. Table 3) in accordance with D'Alessandro & Scheer's (2015) seminal proposals for a modular PIC. Clearly, the finer details of the cross-dialectal patterns outlined here interpreted in terms of the differing parametric mappings of phasal domains at the interfaces remain to be worked out, but they certainly provide further substantial support for D'Alessandro & Scheer's modular PIC which offers a highly promising way forward towards understanding the structural regularities that underlie what might otherwise be written off as superficial phonological irregularities.

References

- Acquaviva, Paolo. 1989. Aspetti della complementazione frasale. Degree thesis, University of Pisa.
- Andalò, Adrina. 1991. Il raddoppiamento sintattico nel dialetto di Napoli. In Luciano Giannelli (ed.), *Tra rinascimento e strutture attuali. Saggi di linguistica italiana*, 241–251. Turin: Rosenberg & Sellier.
- 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, 16–51. Oxford: Oxford University of Press.
- Belletti, Adriana. 2005. Extended doubling and the VP periphery. *Probus* 17. 1–35. <https://doi.org/10.1515/prbs.2005.17.1.1>
- Biberauer, Theresa & Roberta D'Alessandro. 2006. Syntactic doubling and the encoding of voice in eastern Abruzzese. In Donald Baumer, David Montero & Michael Scanlon (eds.), *Proceedings of the 25th West Coast Conference on Formal Linguistics*, 87–95. Somerville, MA: Cascadilla Proceedings Project.
- Bonami, Olivier, Gilles Boyé & Jesse Tseng. 2014. An integrated approach to French liaison. In Paola Monachesi, Gerhard Jäger, Gerald Penn & Shuly Wintner (eds.), *Proceedings of Formal Grammar 2004*, 29–45. Stanford: CSLI Publications.
- Bonet, Eulàlia, Lisa Cheng, Laura Downing & Joan Mascaró. 2019. (In)direct reference in the phonology-syntax interface under phase theory: A response to modular PIC (D'Alessandro and Scheer 2015). *Linguistic Inquiry* 50. 751–777. https://doi.org/10.1162/ling_a_00324
- Bošković, Željko. 1995. Case properties of clauses and the greed principle. *Studia Linguistica* 49. 32–53. <https://doi.org/10.1111/j.1467-9582.1995.tb00464.x>
- 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 Otero & 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>
- Chomsky, Noam. 2012. Foreword. In Ángel Gallego (ed.), *Phases: Developing the framework*, 1–9. Berlin: de Gruyter. <https://doi.org/10.1515/9783110264104.1>
- Cinque, Guglielmo. 1999. *Adverbs and functional heads. A cross-linguistic perspective*. Oxford: Oxford University Press.
- Corr, Alice. 2017. Ibero-Romance and the syntax of the utterance. Doctoral dissertation, University of Cambridge.
- D'Alessandro, Roberta & Tobias Scheer. 2015. Modular PIC. *Linguistic Inquiry* 46(4). 593–624. https://doi.org/10.1162/LING_a_00195
- D'Imperio, Mariapaola, Gorka Elordieta, Sonia Frota, Pilar Prieto & Marina Vigário. 2005. Intonational phrasing in Romance: The role of syntactic and prosodic structure. In Sonia Frota, Marina Vigário & Maria João Freitas (eds.), *Prosodies*, 59–98. The Hague: Mouton de Gruyter.
- D'Ovidio, Francesco. 1874. Fonetica del dialetto di Campobasso. *Archivio Glottologico Italiano* 4(1). 145–184.
- De Blasi, Nicola & Luigi Imperatore. 2000. *Il napoletano parlato e scritto. Con note di grammatica storica (Nuova edizione)*. Naples: Dante & Descartes.
- Diesing, Molly. 1992. *Indefinites*. Cambridge, MA: MIT Press.

- Durand, Jacques & Chantal Lyche. 2008. French liaison in the light of corpus data. *Journal of French Language Studies* 18. 33–66. <https://doi.org/10.1017/S0959269507003158>
- Elordieta, Gorka, Sonia Frota, Pilar Prieto & Marina Vigário. 2003. Effects of constituent length and syntactic branching on intonational phrasing in Ibero-Romance. In Maria-Josep Solé, Daniel Recasens & Joaquín Romero (eds.), *Proceedings of the 15th International Congress of Phonetic Sciences*, 487–490. Barcelona: Causal Productions.
- Fanciullo, Franco. 1986. Syntactic reduplication and the Italian dialects of the centre-south. *Journal of Italian Linguistics* 8. 67–103.
- Fanciullo, Franco. 1997. *Raddoppiamento sintattico e ricostruzione linguistica nel sud italiano*. Pisa: ETS.
- Gallego, Ángel. 2010. *Phase theory*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.152>
- Givón, Talmy. 1990. *Syntax. A functional-typological introduction*, Volume 2. Amsterdam: John Benjamins. <https://doi.org/10.1075/z.syn2>
- Graffi, Giorgio. 1996. Alcune riflessioni sugli imperativi italiani. In Paola Benincà, Guglielmo Cinque, Tullio De Mauro & Nigel Vincent (eds.), *Italiano e dialetto nel tempo. Saggi di grammatica per Giulio C. Lepschy*, 133–148. Rome: Bulzoni.
- Inkelas, Sharon & Draga Zec. 1990. *The phonology-syntax connection*. Chicago: Chicago University Press.
- Kaisse, Ellen. 1985. *Connected speech: The interaction of syntax and phonology*. Orlando: Academic Press.
- Kayne, Richard. 1991. Romance clitics, verb movement and PRO. *Linguistic Inquiry* 22. 647–686.
- Kayne, Richard. 1994. *The antisymmetry of syntax*. Cambridge, MA: MIT Press.
- Ledgeway, Adam. 2000. *A Comparative syntax of the dialects of southern Italy: A minimalist approach*. Oxford: Wiley-Blackwell.
- Ledgeway, Adam. 2009a. *Grammatica diacronica del napoletano* (Beihefte zur Zeitschrift für romanische Philologie Band 350). Tübingen: Max Niemeyer. <https://doi.org/10.1515/9783484971288>
- Ledgeway, Adam. 2009b. Aspetti della sintassi della periferia sinistra del Cosentino. In Diego Pescarini (ed.), *Studi sui dialetti della Calabria* (Quaderni di lavoro ASIt 9), 3–24. Padua: Unipress.
- Ledgeway, Adam. 2016. The dialects of southern Italy. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford guide to the Romance languages*, 246–269. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199677108.003.0016>
- Ledgeway, Adam. 2018. Phonological correlates of syntactic structure: The distribution of *raddoppiamento fonosintattico* in Calabrian. In Mirko Grimaldi, Rosangela Lai, Ludovico Franco & Benedetta Baldi (eds.), *Structuring variation in Romance linguistics and beyond: In honour of Leonardo M. Savoia*, 283–296. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.252.19led>
- Ledgeway, Adam. 2019. Parameters in the development of Romance perfective auxiliary selection. In Michela Cennamo & Claudia Fabrizio *Historical linguistics 2015: Selected papers from the 22nd International Conference on Historical Linguistics, Naples, 27–31 July 2015*, 343–384. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.348.17led>
- Ledgeway, Adam. 2020. The causative construction in the dialects of southern Italy and the phonology-syntax interface. In Giulia Bellucci, Ludovico Franco & Paolo Lorusso (eds.), *Linguistic variation: Structure and interpretation*, 371–400. Berlin: de Gruyter.
- Ledgeway, Adam & Alessandra Lombardi. 2005. Verb movement, adverbs and clitic positions in Romance. *Probus* 17. 79–113. <https://doi.org/10.1515/prbs.2005.17.1.79>

- Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman & Sandra Thompson (eds.), *Clause combining in grammar and discourse*, 181–225. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.18.09leh>
- Loporcaro, Michele. 1988. History and geography of *raddoppiamento fonosintattico*: Remarks on the evolution of a phonological rule. In Pier Marco Bertinetto & Michele Loporcaro (eds.), *Certamen Phonologicum*, 341–387. Turin: Rosenberg & Sellier.
- Loporcaro, Michele. 1997. Lengthening and *raddoppiamento fonosintattico*. In Martin Maiden & Mair Parry (eds.), *The dialects of Italy*, 41–51. London: Routledge.
- Loporcaro, Michele. 1998. L'assimilazione fonosintattica di -t finale nei dialetti della zona Lausberg. In Luciano Agostiniani, Maria Giovanna Arcamone, Fiorella Imparati, Onofrio Carruba & Riccardo Rizza (eds.), *Do-ra-qu pe-re. Studi in memoria di Adriana Quattordio Moreschini*, 237–244. Rome: IEPI.
- Maiden, Martin. 1995. *A linguistic history of Italian*. London: Longman.
- Manzini, Maria Rita & Leonardo Savoia. 2005. *I dialetti italiani e romanci. Morfosintassi generativa* (3 volumes). Alessandria: Edizioni dell'Orso.
- Manzini, Maria Rita & Leonardo Savoia. 2016. Quale informazione sintattica viene interpretata dai fenomeni fonosintattici? Evidenza dalle varietà italiane meridionali. In Adam Ledgeway, Michela Cennamo & Guido Mensching (eds.), *Actes du XXVII^e Congrès international de linguistique et de philologie romanes* (Nancy, 15–20 juillet 2013). Section 4: Syntaxe, 403–414. Nancy: ATILF. <http://www.atilf.fr/cilpr2013/actes/section-4.html> (10 May, 2021)
- Massam, Diane. 1985. Case theory and the projection principle. Doctoral dissertation, MIT.
- Masutti, Vania. 2016. La liaison entre phonologie, morphologie et syntaxe. In Adam Ledgeway, Michela Cennamo & Guido Mensching (eds.), *Actes du XXVII^e Congrès international de linguistique et de philologie romanes* (Nancy, 15–20 juillet 2013). Section 4: Syntaxe, 291–302. Nancy: ATILF. <http://www.atilf.fr/cilpr2013/actes/section-4.html> (10 May, 2021)
- Morin, Yves-Charles & Jonathan Kaye. 1982. The syntactic bases for French liaison. *Journal of Linguistics* 18. 291–330. <https://doi.org/10.1017/S0022226700013621>
- Munaro, Nicola. 2004. Computational puzzles of conditional clause preposing. In Anna-Maria Di Sciullo & Rodolfo Delmonte (eds.), *UG and external systems*, 73–94. Amsterdam: John Benjamins.
- Munaro, Nicola. 2010. Towards a hierarchy of clause types. In Paola Benincà & Nicola Munaro (eds.), *Mapping the left periphery*, 125–162. Oxford: Oxford University Press.
- Plann, Susan. 1986. On case-marking clauses in Spanish: Evidence against the case resistance principle. *Linguistic Inquiry* 17. 336–345.
- Rao, Rajiv. 2008. Observations on the roles of prosody and syntax in the phonological phrasing of Barcelona Spanish. *The Linguistics Journal* 3. 85–131.
- Raposo, Eduardo. 1987. Case theory and Infl-to-Comp: The inflected infinitive in European Portuguese. *Linguistic Inquiry* 18. 85–110.
- Richards, Mark. 2007. On feature inheritance: An argument from the phase impenetrability condition. *Linguistic Inquiry* 38. 563–572. <https://doi.org/10.1162/ling.2007.38.3.563>
- Rivero, María-Luisa. 1994a. Negation, imperatives and Wackernagel effects. *Rivista di linguistica* 6. 39–66.
- Rivero, María-Luisa. 1994b. Clause structure and V-movement in the languages of the Balkans. *Natural Language & Linguistic Theory* 12. 63–120. <https://doi.org/10.1007/BF00992746>
- Rivero, María-Luisa & Arhonto Terzi. 1995. Imperatives, V-movement and logical mood. *Journal of Linguistics* 31. 301–332. <https://doi.org/10.1017/S0022226700015620>

- Rizzi, Luigi & Leonardo Savoia. 1993. Conditions on /u/ propagation in southern Italian dialects. In Adriana Belletti (ed.), *Syntactic theory and the dialects of Italy*, 252–318. Turin: Rosenberg & Sellier.
- Roberts, Ian. 2005. *Principles and parameters in VSO languages*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195168211.001.0001>
- Rohlf, Gerhard. 1966. *Grammatica della lingua italiana e dei suoi dialetti, Volume 1: Fonetica*. Turin: Einaudi.
- Rohlf, Gerhard. 1983. Distinzione di due congunzioni nei dialetti d'Italia (nel senso del latino *ut e quod o quia*). In Paola Benincà, Michele Cortelazzo, Aldo Luigi Prosdocimi, Laura Vanelli & Alberto Zamboni (eds.), *Scritti linguistici in onore di Giovan Battista Pellegrini*, 147–154. Pisa: Pacini.
- Sampson, Rodney. 2016. Sandhi phenomena. In Adam Ledgeway & Martin Maiden (eds.), *The Oxford guide to the Romance languages*, 669–680. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199677108.003.0040>
- Savoia, Leonardo M. 1987. Teoria generativa, modelli fonologici e dialettologia. La propagazione di *u* in una varietà lucana. *Rivista Italiana di Dialettologia* 11. 185–263.
- Savoia, Leonardo M. 2015. *I dialetti italiani. Sistemi e processi fonologici nelle varietà di area italiana e romancia*. Pisa: Pacini.
- Selkirk, Elizabeth. 1974. French liaison and the X' convention. *Linguistic Inquiry* 5. 573–590.
- Selkirk, Elizabeth. 1984. *Phonology and syntax: The relationship between sound and structure*. Cambridge, MA: MIT Press.
- Selkirk, Elizabeth. 1986. On derived domains in sentence phonology. In Colin Ewen & John Anderson (eds.), *Phonology Yearbook* 3, 371–405. Cambridge: Cambridge University Press. <https://doi.org/10.1017/S0952675700000695>
- Selkirk, Elizabeth & Koichi Tateishi. 1988. Syntax and phonological phrasing in Japanese. In Carol Georgopoulos & Roberta Ishihara (eds.), *Interdisciplinary Approaches to Language: Essays in honor of S.Y. Kuroda*, 316–336. Dordrecht: Reidel.
- Sheehan, Michelle. 2014. Portuguese, Russian and the theory of control. In Hsin-Lun Huang, Ethan Poole & Amanda Rysling (eds.), *Proceedings of the 43rd Annual Meeting of the North East Linguistic Society 43 (NELS 43)*, Volume 2, 115–126. Amherst, MA: GLSA.
- Sheehan, Michelle. 2018a. Control of inflected infinitives in European Portuguese. In Anabela Gonçalves & Ana Lúcia Santos (eds.), *Complement clauses in Portuguese: Adult syntax and acquisition*, 27–58. Amsterdam: Benjamins. <https://doi.org/10.1075/ihll.17.02she>
- Sheehan, Michelle. 2018b. On the difference between exhaustive and partial control. In Federica Cognola & Jan Casalicchio (eds.), *Understanding null subjects: A synchronic and diachronic perspective*, 141–170. Oxford: Oxford University Press. <https://doi.org/10.1093/oso/9780198815853.003.0006>
- Silvestri, Giuseppina. 2007. Il rafforzamento fonosintattico nel dialetto di Verbicaro (Cosenza). *L'Italia Dialettale* 68. 146–151.
- Silvestri, Giuseppina. 2009. La metafonía nel dialetto di Verbicaro (CS). *Italia Dialettale* 70. 169–226.
- Silvestri, Giuseppina. 2014. *Rafforzamento fonosintattico* as a phenomenon of the phonology-syntax interface. Paper presented at the Romance Linguistics Seminar, University of Cambridge, October 28.
- Stowell, Tim. 1981. Origins of phrase structure. Doctoral dissertation, MIT.

- Torcolacci, Giuseppe. 2014a. Il raddoppiamento fonosintattico e la codifica di tratti morfosintattici. Il caso dei dialetti italiani meridionali. In Roberta D'Alessandro, Claudio Di Felice, Irene Franco & Adam Ledgeway (eds.), *Approcci diversi alla dialettologia italiana contemporanea. Special issue of L'Italia dialettale* 75. 247–271.
- Torcolacci, Giuseppe. 2014b. *Marking the default. Auxiliary selection in southern Italian dialects*. Leiden: LOT.
- Torrego, Esther. 1998. *The dependencies of objects*. Cambridge, MA: MIT Press.
<https://doi.org/10.7551/mitpress/2337.001.0001>
- Tortora, Christina. 2014. *A comparative grammar of Borgomanerese*. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199945627.001.0001>
- Truckenbrodt, Hubert. 1999. On the relation between syntactic phrases and phonological phrases. *Linguistic Inquiry* 30. 219–255. <https://doi.org/10.1162/002438999554048>
- Uriagereka, Juan. 2011. Derivational cycles. In Cedric Boeckx (ed.), *The Oxford handbook of minimalism*, 239–259. Oxford: Oxford University Press.
<https://doi.org/10.1093/oxfordhb/9780199549368.013.0011>
- Zanutini, Raffaella. 1997. *Syntactic properties of sentential negation. A comparative study of Romance languages*. Oxford: Oxford University Press.

The causative-inchoative alternation (as we know it) might fall short

Crosslanguage systematicities and untapped data
from Romance and Greek

M. Eugenia Mangialavori Rasia

National Research Council-Argentina (CONICET)

This contribution challenges the claim that internal arguments are stable/constant arguments in the causative-inchoative alternation. It presents evidence that a third (external-argument-only) variant, produced by the standard combinatorial system, is possible (and systematic) in Romance and Greek. Free composition with a null causative v^0 independent of the internal-argument-licensing head: explains all the hallmarks of monadic (intransitive) variants; correctly preserves event/argument structure correlation (no internal-argument-introducing head, no change-of-state event); uncovers crosslanguage contrasts concerning \pm availability of cause(r) interpretation of sole arguments in equipollent derivations. My argument is supported by a parallel with non-Romance languages with comparable morphology (Greek). Such symmetries show a transparent morpho-semantic-syntactic correlation in the choice of argument frame, extending to other verb classes with transitivity alternation. Greek and Romance data support a (a) wider causative alternation with expected semantic/syntactic/morphological implications; (b) (missing) structural distinction among (in)transitivity alternations. Arguably, languages may differ in the availability of this option, showing at least two patterns of variation.

Keywords: argument structure, alternation, intransitivity, causativity, stativity, syntax-semantics interface, Romance, Greek, changing morphology, VP-internal composition

1. Introduction

The causative-inchoative alternation, illustrated in (1), has stirred great interest in the literature. The reason, in part, is the suggestive correlation noted between syntax (argument realization) and semantics (event composition) (2).

- (1) a. *Persistent acid rains corroded the pipes.* (Causative)
 b. *The pipes corroded.* (Inchoative)
- (2) a. [_{CAUSE} *persistent acid rains* [_{BECOME} [*the pipes*<_{STATE}*corroded*>]]]
 b. [_{BECOME} [*the pipes*<_{STATE}*corroded*>]] (Levin & Rappaport-Hovav 1995: 85)

In the analysis, a common assumption is that causatives involve complex semantic structures where two subevents ('cause', 'process') combine in an 'implicational' (Hale & Keyser 2002) relation. A fundamental rule of event composition underlies this claim: 'cause', if present, causally implicates 'process' (3). Similarly, it is claimed that the causative form of these verbs encompasses the simpler structure corresponding to the intransitive variant (4) (Levin & Rappaport-Hovav 1995; Hale & Keyser 2002).

The way in which these two variants obtain – and, subsequently, which is the correct phrase structure of these verbs –, however, has been the subject of much debate. Some researchers argue for fully-specified lexical representations, generally defending a basic transitive frame, from which the inchoative variant derives (Levin & Rappaport-Hovav 1995; Reinhart 2002; Chierchia 2004); others instead see the structural complexity of the transitive variant as the consequence of syntactic composition, assuming a basic unaccusative configuration instead (in works from different theoretical orientations, e.g., Dowty 1979; Hale & Keyser 1993; Levin & Rappaport-Hovav 2012). More recently, 'nonderivational' approaches (Rosen 1996; Piñón 2001; Doron 2003; Alexiadou et al. 2006) have appeared, proposing instead that neither variant derives from the other (e.g., by a lexical rule). Rather, two distinct variants are constructed from the same source.

- (3) $e_1 \rightarrow e_2$ (Hale & Keyser 1993: 69, 1988: 46)
- (4) $V_{\text{TRANSITIVE}=\text{CAUSE}} \text{ to } V_{\text{INTRANSITIVE}}$ (Levin 1993: 27)

Certain empirical observations are, nonetheless, central to the discussion. In a variety of languages – including those under discussion here (Haspelmath 1993, among others) – inchoative (unaccusative) variants generally take special morphology (Romance: *se/si*-cliticization). While some accounts see this as a sign of further derivation (i.e., in support of the 'derived' status of the inchoative; Piñón 2001; Alexiadou et al. 2006 for discussion), others propose instead that changing morphology in the alternation is a visible reflex of a choice of light verb (Harley 2012). It is the latter view, I suggest, that better captures Romance facts.

I follow the nonderivational theoretical angle set out in Folli & Harley (2005),¹ while at the same time building on a central premise from fundamental works on

1. In my view, a nonderivational account does not pose a conflict with presenting the question in the frame of the causative alternation – at least from the perspective, adopted here, that scs instantiate an alternate argument structure realization, and, more importantly, one involving the same arguments the Causative Alternation crucially bears on.

lexical syntax (Hale & Keyser 1993); concretely, on Hale & Keyser's account, to say that a verb participates in the causative-inchoative alternation means that an independent 'notional type of V', which is a 'dynamic event' (Hale & Keyser 1993: 71) (represented as V_2 in the configuration), can freely appear as the complement of the basic monadic configuration yielded by v^0 (V_1 in (5a)).

- (5) Internal composition of alternating verbs
- a. [V_1 [V_2 , DP [V_2^0 , A]] (Hale & Keyser 2002)
 - b. [X CAUSE [y BECOME <state>]] (Levin & Rappaport-Hovav 1995)
 - c. [DP, *Init* [DP, *ProcP* (ResP)]] (Ramchand 2008)

I take the hypothesis that neither the capacity of licensing an internal argument nor eventivity are default properties of verbs typically instantiating the alternation, on the assumption that much is to be learned if one tests the limits of this hypothesis. The subsequent aim of this paper is to investigate the true default semantic and syntactic properties of verbs showing causative alternation, thus raising the question as to what really belongs to the verb's lexical entry and what is compositionally (optionally) added in the course of the derivation – and, ultimately, how languages may differ in this respect.

To do this, I draw attention to the idea that the causative-inchoative alternation in various languages, including most Romance varieties, extends to a logically plausible third frame (6c) that the combinatorial system proposed above makes possible, but whose existence is not commonly addressed in the literature.

- (6) a. *Estos minerales corroen algunas tuberías.*
 these minerals corrode some pipes
 "These minerals corrode some pipes."
 b. *Algunas tuberías se corroen.*
 some pipes INCH corrode
 "Some pipes corrode."
 c. *Estos minerales corroen.*
 these minerals corrode
 "These minerals are corrosive."
 (cf. English "These minerals become corroded.")

These variants pose interesting challenges for various generalizations on change-of-state verbs and the causative alternation. Notably, by combining cause(r) interpretation of the sole DP with stative behavior (shown below), they become problematic for the event composition rule (3), according to which, 'cause', if present, implicates 'process'. Stative-Causatives [SC] variants are also a problem for the claim that 'cause(r)' interpretation in alternating verbs is purely structural (an automatic result of transitivity, obtained by merging the independent internal-argument-licensing V^0 [V_2] with the monadic external-argument-introducer categorizer [V_1 in (5)], as Hale & Keyser 2002: 176; Zubizarreta & Oh 2007, among others,

propose). Additionally, SCs challenge the generalization that unique arguments of so-called change-of-state verbs are by default direct internal arguments (Levin & Rappaport-Hovav 2010: 293). Conversely, they instantiate the Immediate Cause Linking Rule:

- (7) The argument of a verb that denotes the immediate cause of the eventuality described by that verb is its external argument.

(Levin & Rappaport-Hovav 1995)

Drawing on previous constructionalist accounts of the causative alternation (§ 2),² I propose an account for SCs based on a null causative v^0 freely available for derivation. If correct, direct composition with the causative v^0 – as a systematic alternative in Romance³ – would produce the expected results from a configuration crucially missing the internal-argument-licensing head (i.e., the independent V instantiating a ‘dynamic event’, Hale & Keyser 1993: 71; the semantic change-of-state relation, Hale & Keyser 2002: 176; BECOME in Levin & Rappaport-Hovav 1995; the ‘constructed’ BECOME event in Rothstein 2012: 87; V^0_{BECOME} for Folli & Harley 2005, or Ramchand’s 2008 *Proc* in (5c) above). This would produce a structure semantically and syntactically simpler than the causative (dyadic) normally considered (i.e., (5a)), and, at the same time, maximally distinct from the monadic (unaccusative/inchoative) structure commonly analyzed (6b).⁴

Moreover, the alternative would account for a consequent crosslanguage contrast, bearing on the (un)availability of ‘cause(r)’ interpretation of the unique DP in string-identical frames, correlating with the (un)availability of free/systematic SC derivation, as the glosses suggest.⁵

In addition, the proposed analysis correctly predicts systematic production of SCs in further verb classes associated to the causative alternation – namely, psych verbs compatible with subject-cause(r) interpretation – for Romance. This contrasts with the dyadic form seen in languages like English or German.

The discussion (§ 3.1) offers new evidence pointing to the nondefective status of dynamic event, but also to the consequent eventlessness of the causative

2. Crucially, process-less compounding with the causative verbalizer has been used in English to account for Stative-Causative predications (Ramchand 2008), and direct compounding with V_{INIT} , for atransitivity (McIntyre 2004) – that is, two issues central to an account of SCs.

3. Setting aside French (severely restricted) and few other varieties.

4. Hence avoiding a model of partial projection of lexical semantic structure as well as lexical entry proliferation.

5. Most notably, the default undergoer interpretation of the argument in the English literal gloss *Minerals corrode* (=minerals become corroded) vs natural cause reading in Spanish (6c) (=minerals cause corrosion).

(external-argument-introducing)⁶ head in the alternation – as proposed by Ramchand (2008), among others. If correct, Romance (and Greek) could provide empirical evidence for the independence of v_{CAUSE} and of its status as verb-creating categorizing head (Harley 2014, among others).

- (8) a. *El corte de Juan {molesta/irrita/asusta}*. (Spanish)
 the cut of Juan annoys irritates frightens
 “John’s haircut is annoying/irritating/frightening.”
 b. *John’s haircut {annoys/irritates/frightens} *(Rita)*. (English)

We will also touch on some of the key points from non-Romance languages which are of particular relevance to these claims (§ 3.2). Greek shows a surprising morpho-syntactic-semantic correlation, with split morphology (Act/NAct) acting as a visible reflex of a free but nontrivial derivational option, crucially paralleling Romance *se*-marking.

2. A basic monadic eventless configuration

2.1 Why not a null object?

Nonovert arguments (implicit/null/generic arguments, *pro*, A/A'-traces) constructions seem a straightforward answer to the problem for different reasons. First, they share various properties with SCs; notably, the restriction to generic tenses (Rizzi 2003).⁷ Second, on a null object [NO] account, it would be possible to maintain the long-standing notion that internal arguments are constant arguments in the alternation (Hale & Keyser 2002).⁸

Two problems, however, arise. On the one hand, null/unspecified/unexpressed objects are generally disallowed with change-of-state verbs (Levin & Rappaport-Hovav 1995, 2010), while SCs are fine. On the other hand, and more crucially, SCs do not behave as NOs.

6. Assuming, namely, a Voice+v bundling schema for Spanish/Italian (see Folli & Harley 2007; Harley 2017).

7. An alternative analysis would suggest that SCs involve a generic operator. This possibility is readily dealt with without assuming a transitive configuration under Hale & Keyser’s (1998: 48) approach.

8. It would also preserve the analysis of alternating verb as basic transitives, coherent with the unmarkedness (lack of *se*-cliticization) observed.

Namely, while NOs can appear as the subject of a small clause with an adjectival or participial predicate, SCs do not.

(9) NO

- a. *Este horno admite artificiales y naturales.*
 this oven admits artificial.PL and natural.PL
 “This oven admits artificial and natural [items].”

SC

- b. *Este horno calienta (*artificiales/*naturales).*
 this oven heats artificial.PL natural.PL
 (Intended) “This oven has heating capacity (artificial/natural)”

Similarly, Romance transitive verbs generally allow both NO-oriented depictives (10a) and resultatives (10b), also in contrast to SCs (11).

(10) NO

- a. *Este chef compra {empaquetado/natural/orgánico}.*
 this chef buys wrapped natural organic
 “This chef buys wrapped/natural/organic [produce].”
- b. *Este chef cocina {abundante/rico/salado}.*
 this chef cooks abundant tasty salty
 “This chef cooks abundant/tasty/salty [meals]”

(11) SC

- a. *Estos hornos calientan/ asan (*empaquetados⁸/*naturales/*orgánicos).*
 these ovens heat roast wrapped natural organic [items]
 (intended) “These ovens have heating/roasting capacity (wrapped/natural/organic)”
- b. *Este horno calienta/ asa (*abundante/*rico/*saludable).*
 this oven heats roasts abundant tasty/ healthy
 (intended) “This oven has heating/roasting capacity (abundant/tasty/healthy meals)”

While better-known NO tests show that implicit arguments bind participial/adjectival predicates (Massam 1989), as in (12), (13) extends the contrast seen above ((9)–(11)).

(12) NO (Rizzi 2003)

(Italian)

- Un dottore serio visita Ø_i nudi_i.*
 “A serious doctor visits nude.”

9. (11a) is possible, but on a subject-oriented reading of the predicate, paraphrasable as “Wrapped ovens have heating capacity” (vs the NO default reading obtained in English (\approx *These ovens heat wrapped* [\emptyset])).

- (13) SC
- a. *El fármaco cura (*enfermos/*afebrados).* (Spanish)
(Intended) “The drug has healing capacity on sick/feverish.”
- b. *Un buon farmaco cura *(Ø_i ammalati).* (Italian)
(Intended) “A quality drug has healing capacity on sick.”

Unlike object-licensing frames – implicit arguments (cf. Bhatt & Pancheva 2006; Alexiadou et al. 2006) and passives included – , SCs do not allow PRO-control.¹⁰

- (14) Transitive/Passive
- a. *Estos hornos los calientan/asan /tuestan (para ser vendidos).*
“These ovens heat/roast/toast them (to be sold away).”
- b. *Se calientan/asan/tuestan (para ser vendidos).*
“They are heated/roasted/toasted (to be sold away).”

- (15) SC
*Estos hornos calientan/asan/tuestan *(para ser vendidos).*
(Intended) “These ovens have heating/roasting/toasting capacity to be sold.”

As opposed to NOs (Rizzi 2003, among others), SCs (16) do not allow NO quantification (also (22) below) and fail to bind reflexive pronouns.¹¹

- (16) *El sol calienta (*todos/ *algunos/ *a sí mismo).*
the sun warms all some itself
“The sun instigates heat (causes heat) (*all/*some/*to itself).”
- (17) a. *La injusticia {entristece/enoja} (*consigo mismo).* (SC)
the injustice saddens maddens with himself
“Injustice saddens/maddens (*with oneself).”
- b. **{se/lo} {entristece/enfada} (consigo mismo).*
INCH/ACC saddens upset with-self same
“He gets (him) sad/mad at himself” (Inchoative/Transitive)

(17) also suggests that SCs are systematically productive with unpassivizable verbs and verbs usually not allowing NOs (Cf. (18)) – nor object alternations altogether (Levin 1993). (19) illustrates what is a highly productive construction in Spanish – and, apparently, in Romance.

10. Like in (11), the construction could be admitted under external-argument-targeted interpretation (*These ovens heat so they get sold* (odd)). Conversely, an inchoative reading of *se*-cliticization (odd depending on subject selection) allows PRO-control (“*They get heated /roasted (to be sold away)*”).

11. Cf. (17) vs Rizzi’s Italian example *La buona musica riconcilia Ø con se stessi*. “Good music reconciles Ø with oneself.” Apparently, Italian is less systematic, while Spanish seems consistently restrictive. Still, observations are in order about Rizzi’s Example (Mangialavori Rasia 2019).

- (18) **Juan llenó* Ø. (Spanish)
 **Juan riempì* Ø. (Italian)
 **Juan filled* Ø. (English)
- (19) a. *El pan llena (bastante)*. (Spanish)
Il pane riempe (abbastanza). (Italian)
 **The bread fills (enough)*. (English)
 “Bread is (quite) filling.”
- b. *Este producto suaviza*. (Spanish)
Questo prodotto ammorbidisce. (Italian)
 **This product softens*. (English)
 “The product is softening/has softening properties.”
- c. *La música clásica relaja*. (Spanish)
La musica classica rilassa. (Italian)
 *(*The*) *classical music relaxes*. (English)
 “Classical music is relaxing.”

Finally, Italian data is key to show that *ne*-cliticization – tightly connected with NOs (Russi 2008: 113) –, which is another pattern expected in NOs, is also disallowed in SCs.¹²

- (20) *La radiazione infrarossa (ne) riscalda/brucia* *(*la metà*).
 the radiation infrared of-them heats burns (the half)
 “Infrared radiation burns *(half) (of them).” (NO)
 “Infrared radiation has heating/burning capacity (*half (of them)).” (SC)

2.2 Eventless monoargumental (and eventless cause)

Above, different facts suggest that the object is missing in a more radical sense (total absence of structure). Interestingly, aspectual tests give results fully consistent with the processless predicate expected from a structurally simpler configuration.

First, in contrast to internal-argument-licensing frames, SCs are invariably atelic (21).

- (21) *El grafito *(lo/se) calienta (hasta quedar incandescente/fluido)*.
 the graphite ACC/INCH warms until remain.INF incandescent fluid
 “Graphite heats (**it/up**) (until incandescent/fluid).” (Inchoative)
 (Intended) “Graphite has heating capacity (until incandescent/fluid)” (SC)

12. Expected inasmuch as a focal post-V quantificational expression is missing.

Here, the correlation between inchoative/accusative morphology, and the capacity to license modifiers describing endpoint/result of a change-of-state event, is consistent with a defective eventless predication in absence of the relevant V head – associated, recall, to a ‘dynamic’ event in the literature. If correct, the contrast between SCs and inchoatives coincides with an opposition between ‘Cause’ vs ‘Undergoer’ interpretation of the sole argument – *el grafito* “the graphite” in (21) – in two radically different monoargumental variants. Crucially, these variants are told apart by morphological marking. Interestingly also, the distinction further correlates with a contrast between invariably atelic (processless) eventualities vs proper (change-of-state) events with some conceptually-identifiable result in a participant (in principle, the state in (5b)). Notice that SCs failing to derive resultative readings is aspectually consistent with (11) above.¹³

Assuming that bare quantifiers behave as NOs (Cattaneo 2008), perfective inflection and culmination also suggest that eventivity and (variable) telicity are constrained to internal-argument-licensing structures (cf. *#Questo ha infuriato.*), hence pointing to the possibility that these properties are not part of the verb’s inherent (lexical) denotation, but rather the result of composition with the internal-argument-licensing V°.

- (22) *Questo ha {infuriato/impoverito/indignato} # (molti).* (Italian)
Esto ha {enfurecido/empobrecido/indignado} # (a muchos). (Spanish)
 “This has infuriated/impoverished/outraged (many)” (Transitive/NO)
 “This is infuriating/impoverishing/outraging #many” (SC)
- (23) *El sol quema (en un minuto/completamente) # (algo/alguno).* (Spanish)
 “The sun burns some[thing]/some [things] in a minute/completely”
 (Inchoative)
 “The sun has burning capacity #in a minute/#completely” (SC)

In consequence, telicity (tracked by *casi* “almost”), progression, and perfectivity are possible only if a unique DP is interpreted as undergoer.¹⁴ The key observation here is that perfective/telic senses force undergoer (inchoative) reading in otherwise string-ambiguous contexts (verbs with morphologically unmarked inchoatives). In SCs, there is no process, and there is no (affected) entity on which to estimate event progression either.

13. Insofar as the symmetry between lack of result state and of lack of telicity is concerned.

14. A comparable structural ambiguity is visible in English *be+ing* (para)phrases, giving both SC (property of the subject (present participle used as attributive adjectival)) and change-of-state (progressive, undergoer DP) readings (cf. *This chocolate is fattening* ≈ causes fatness/is getting fat). For some reason, it is a canonical ILP (copula-headed) predication that is recruited as semantic analogue for SCs when unavailable.

- (24) #*El chocolate casi engorda/está engordando/engordó.*
 “Chocolate almost gets fat/is fattening/fattened”

In fact, SCs show patterns expected in ‘eventless’ atelic types (states). Notably, imperatives (Jackendoff 1990, among others) are disallowed.

- (25) NO
Calienta a fuego fuerte, revolviendo.
 heat Ø at strong fire stirring
 “Heat (up) on high fire, stirring.”
- (26) SC
- a. *Horno, calienta!*
 (intended) “Oven, have heating capacity!”
 [*lit.* “Oven, warm!”]
 - b. *El horno calienta (#a fuego fuerte).*
 (intended) “The oven has heating capacity (at high flame)”
 [*lit.* “The oven warms at strong fire.”]

SCs are also incompatible with manner modification (26b), as opposed to internal-argument-licensing variants, including NOs (25) and middles (27b) – which allegedly involv NOs as well, Massam (1992). As known (Rothmayr 2009), it is precisely the event argument that anchors the adverbial to the sentence. Therefore, even if SCs and middles share key properties – ‘stativization’^{15, 16} of the predicate

15. I use the original term in Massam’s description, which does not amount to saying that I entail or agree that SCs are a result of some ‘stativization’ process as well (or that stativity is derived, for that matter). See fn. below.

16. An anonymous reviewer contends that state verbs normally combining with themes and/or experiencers as subjects is an exceptionless generalization that the present analysis contradicts, while no such problem arises if stativity is seen as the result of a ‘stativization process’.

On the one hand, the present analysis follows the by now widely-embraced view that aspect (stativity, telicity, eventivity) is defined in structural terms (Jackendoff 1991, among others) and not as a property of the ‘lexical’ verb per se, as in the prototypical example *The wall surrounds the city* (Rothmayr 2009: 38, among others). On the other hand, it seems that a ‘stativization process’ would probably involve some kind of structural (lexical? syntactic?) operation that is not represented in the syntax or in the semantic structure, at least according to data. The idea of a non-derived configuration is in consonance with Hale & Keyser’s insight that the external-argument-selecting head is independent of the inchoative-creating V. The version of their approach that we adopt does not require a layering of functional heads (notably, the embedded v head responsible for introducing the eventive component) requiring some kind of further operation afterwards. Additionally, a simpler structure would more easily accommodate the observed morphological and semantic (event-related) simplicity. Inasmuch as causative/resultative structures are generally taken to be complex structures built from simpler ones, a further derivation ([v[V]; [[cause] [become]]; or [v[SC] depending on the account) for eventivity seems the normal course of events. In fact, it has been observed that “it is precisely this possibility that

(Roberts 1987; apud Massam 1992), occurrence in generic tenses, attribution of a property reading to the subject (Hale & Keyser 1988) – the fact that inchoative/passive/impersonal (*se*) and accusative morphology are nontrivial in licensing the adverbial supports the observation that the internal-argument-licensing structure produces relevant minimal pairs. Moreover, these nontrivial pairs reflect a coherent asymmetry in event structure.¹⁷ The syntax/semantics correspondence that links eventivity to the licensing of an internal argument is thus preserved exactly in every case. This can be easily seen by pairing inchoative and a middle variants (both sharing *se*) vs. the cliticless SC.

- (27) a. *El chocolate # (te) engorda fácilmente.*
 ok“Chocolate makes you fat easily.” (Transitive (cliticized))
 #“Chocolate has fattening properties easily.” (SC (no clitic))
- b. *Este horno # (se) calienta (fácilmente).*
 ok“This oven heats (up) easily.” (Inchoative/Middle (cliticized))
 ok“This oven is easily heated up.” (Passive (cliticized))
 #“This oven has heating capacity easily.” (SC (no clitic))

The impossibility of endpoint modification was anticipated above. Yet, note that if *for-x-time* occurs at all in SCs, it is not interpreted as an event variable (nor as a result-state adverbial, Dowty 1979), but as a temporal bound on a more or less permanent property – i.e., an individual-level predicate [ILP]. Significantly, ILPs are considered an eventless (Kimian) stative type (Maienborn 2007).

is exploited by object alternations” (Levin manuscript, among others). The proposal of a core eventivity is also in direct conflict with notions with high currency in the literature; notably, that inherent aspectual properties in a verb are expected to represent components of meaning lexicalized by a particular verb *in all the contexts it is used in* (Levin manuscript, among others).

17. An anonymous reviewer asks how, while the contribution convincingly demonstrates that SCs have generic property readings in keeping with the syntactic absence of the internal argument, one can explain that generic readings arise with suitable overt objects. Yet, it seems that genericity in such cases bears on a universal quantifier (object-defined) (cf. the proposed Example (ib)).

The SC variant denotes a property of ovens in general, not a universal capacity on items in a token (a specific oven, over generic stuff). Note that a modifier like *again* gives the classic ambiguity expected in resultatives for (b) examples, while, if allowed in SCs, it would only give a reading in which the (individual-level) property is somehow reestablished (cf. Spanish *El horno calentó todo nuevamente* [lit. “The oven heated everything again”] vs *El horno #calentó nuevamente* “The oven would have heating capacity again” (SC)). This is predicted by our analysis.

- (i) a. *Cuptorul încălzește.* (Romanian)
 oven.the heats
- b. *Cuptorul încălzește orice.*
 oven.the heats anything.

- (28) *El chocolate engorda/ los payasos asustan (#en/??por un tiempo).*
 the chocolate fattens the clowns frighten in/for a time
 “Chocolate is fattening/Clowns are frightening for a while (=not all their lives).”

Moreover, modifiers like *a little* (Rothmayr 2009, among others) also give stative (degree) readings – as opposed to the event-related reading in the inchoative (e.g., *Acids corrode for a while*). Present tense in SCs does not license habitual or frequentative readings either, thereby contrasting with ‘eventive’ atelic predicates.

- (29) *Los ácidos corroen mucho/un poco.* (Spanish)
Gli acidi corrodono molto/un po’. (Italian)
 “Acids are highly/a little corrosive.”

Finally, SCs do not allow anaphoric reference by *this happened ...* (30),¹⁸ perception reports (31), spatial location, manner adverbials, instrumentals, comitatives (32), and other functionally event-integrated participants (33). This is the behavior expected for eventless stative types (Maienborn 2007).

- (30) *Los payasos #(se) asusta(#ro)n; esto sucedió cuando ...*
 the clowns INCH frighten(ed); this happened when...
 “The clowns got frightened; this happened when ...” (Inchoative (cliticized))
 #“Clowns have been frightening; this happened when ...” (SC (unmarked))
- (31) #*Ví a los payasos asustar.*
 saw ACC the clowns frighten.INF
 #“I saw the clowns being frightening.”
- (32) *Los payasos (se) asustan en sus casas.*
 the clowns INCH frighten in their homes
 “Clowns get frightened at their homes.” (Inchoative (cliticized))
 #“Clowns are frightening at their homes” (SC (unmarked))
- (33) *Los payasos asustan (#con un palo/ #con amigos/ #voluntariamente/ #metódicamente).*
 the clowns frighten with a stick with friends voluntarily
 methodically
 (Intended) “The clowns are frightening (with a stick/with friends/on purpose/ methodically).”

18. Although the test is affected by SC incompatibility with perfectivity, a relevant contrast also obtains in the present (*Los payasos #(se) asustan; esto sucede cuando ...* “Clowns get frightened, this happens when” (Inchoative(cliticized)) vs “Clowns are frightening, #this happens when...” (SC)).

2.3 Section summary

Patterns presented above show how SCs sharply differ from those of sentences commonly analyzed as presenting syntactically represented null objects.

Concomitantly, aspectual tests consistent with a processless predication in SCs show another contrast to both null-object constructions and better-known monadic realizations (inchoatives). In turn, perfective inflection and culmination combine to indicate that eventivity and variable telicity are constrained to internal-argument-licensing structure.

More specific patterns expected in eventless atelic types (states) further indicate that eventivity in these verbs is present only in specific (transitive) frames.

3. Additional evidence

3.1 Psych verbs

Psych verbs (appearing in (30)–(33) above) are relevant to the discussion for various reasons. Besides participating in the alternation (Levin 1993; Alexiadou & Iordăchioaia 2014), psych verbs played an important role in the analysis of stativity, change-of-state and causativity (since Dowty 1979, among others). More importantly, they have been used as examples of ‘Stative-Causative’ constructions in English or German (i.e., in languages where the production of SCs is not free or systematic, e.g., Arad 1998). As a consequence, they have been associated with a configuration headed by the causative head (*Init*^o, v_{CAUSE} in the terminology here adopted). This head produces states under the condition that *Init* does not have a process (V^o) projection in its complement (Ramchand 2008: 64). Such an analysis interests us since it proposes to account for the causative alternation by drawing on independent composition with the external-argument-introducing null causative head (2008: 95).

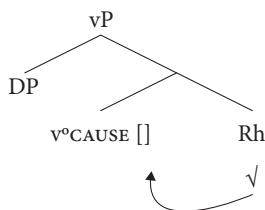
In Spanish, SCs (34c) – fully productive with this verb type – show clear differences with the intransitive middle-like variant traditionally analyzed (34b).

- (34) Middle Alternation (amuse-type verbs, Levin 1993: 191)
- a. *Los payasos asustan/divierten a los niños pequeños.*
“Clowns frighten/amuse little children.”
 - b. *Los niños se asustan/divierten (fácilmente).*
“Little children frighten/amuse easily.”
 - c. *Los payasos asustan/divierten.*
“Clowns are frightening/amusing.”
[lit. “Clowns frighten/amuse.”]

In English, SCs are typically dyadic (Arad 1998) and involve two participants in a conditional relation outside which the caused state “ceases to exist” (Rothmayr 2009:

54). By contrast, in Romance (Spanish, Italian, Catalan, Portuguese, Romanian),¹⁹ the availability of monadic (external-argument-only) frames excludes SCs from this relational condition (recall (8) above). The expression of a cause(r) is thus allowed without requiring the (semantic/syntactic) instantiation of the experiencer. The configuration in (35) should reflect all these points, as well as the key condition on the complement of v_{INIT} (v_{CAUSE}): states produced by the causative head may only host rhematic material as complement.²⁰

(35) Proposed syntax for SCs



A configuration similar to the one proposed by Ramchand has three advantages. It is amenable to the one assigned to unergative verbs in the standard (Hale & Keyser 2002) analysis. It captures the eventlessness of the external-argument-introducing head. And, more importantly, that SCs are freely formed from Object-Experiencer verbs (36a), but not from Subject-Experiencer verbs (36b) – i.e., roots requiring the opposite semantic distribution – follows naturally from the semantic properties of the categorizer ($v_{\text{INIT/CAUSE}}$) in question.²¹

- (36) a. *Katherine {asusta/ irrita/ molesta/ divierte}*.
 Katherine frightens irritates annoys amuses
 “Katherine is frightening/irritating/annoying/amusing.”
- b. **Katherine teme*.
 Katherine fears
 (Intended) “Katherine is fearsome (=causes fear).”

19. Except SCs in, e.g., French, where constructions like *Ça inquiète beaucoup* “This worries a lot” are possible but rare.

20. Ramchand’s suggestions, along with the data presented here, are relevant to a bigger theoretical discussion, notably, the conditions putatively imposed by the head licensing the external argument. There are generalized claims that seem too strong in the face of Romance data here. Namely, that lexical roots of verbs showing causative alternation can only be inserted under $V_{\text{BE-COME}}$ (Folli & Harley 2005, 118). Crucially, direct composition of the root with v_{CAUSE} is argued for, on different data, also by other authors, notably Alexiadou & Anagnostopoulou (2003).

21. A detailed discussion on SC availability for psych verbs, the relevance of the choice of subject, and the asymmetric productivity in Romance as opposed to, e.g., English is offered in Mangialavori & Marin (2021).

3.2 Greek

Greek shows strikingly similar patterns. In verbs with unified morphology (unmarked inchoative), monadic constructions are structurally ambiguous. Either Cause(r) or Undergoer interpretation of the subject are allowed, just like in Spanish.

- (37) *El sol enrojece.* (Spanish)
O ilijos kokinizi. (Greek)
 the sun reddens
 “The sun becomes red.” (Inchoative)
 “The sun causes redness.” (SC)

Otherwise, special morphology correlates with the realization of V_{BECOME} – the V° heading inchoative variants (Folli & Harley 2005, among others) –, reflecting *se*-marking in Romance.

- (38) a. *I supa kegete.* (Spanish: *La sopa se quema*)
 the soup.NOM burn.NAct.3s
 “The soup becomes burnt.”
 b. *I supa kei.* (Spanish: *La sopa quema*)
 the soup.NOM burn.3s
 “The soup causes burns.”

The derivational option thus becomes, for both languages, morphologically visible. Its distribution fits the claim that changing morphology in the causative alternation correlates with a shift from V_{CAUSE} (arguably heading causative variants, recall (35)) to the eventive V_{BECOME} (as key verbal head) yielding the unaccusative monadic variant instead – conforming to the defectiveness of the causative head in Hale & Keyser’s original account. The important point here is that, under a nonderivational account, morphological marking can be naturally seen as the result of a change in choice of light verb, rather than as part of a lexical operation (Harley 2012).²² This perspective, if correct, allows for a fully-integrated analysis in which morphological considerations have clear implications for syntactic and semantic analyses.

Subjects compatible only with cause(r) interpretation restrict the alternation, disallowing inchoative morphology here as well. In the present analysis, four otherwise puzzling characteristics noted in these cases – stative meaning (“inherent property that causes the eventuality expressed by the verb”, Alexiadou & Anagnostopoulou 2003), *pure* state pattern, lack of change of state denotation, and

22. Conversely, in English, where composition with V° seems in general defective (no free choice of verbalizer), V_{BECOME} is also morphologically defective (no changing morphology in the alternation).

unergative behavior (e.g., lack of adjectival passive: **I fotia ine kameni* ‘the fire is burnt’, Alexiadou & Anagnostopoulou 2003) – are simply expected.²³

- (39) *I fotia kei/#kegete.*
 the fire.NOM burns/burns.NAct
 ‘The fire burns (=causes burns).’

(39) further suggests that SCs are freely formed from verbs not allowing NOs (cf. **O Janis ekapse* ‘John burned Ø’),²⁴ yielding consistent eventless patterns – notably, perfective morphology constrained to agentive/undergoer interpretation of *Janis*, i.e., transitive/inchoative variants.

Also here, object realization correlates with eventivity. The alternative creates minimal pairs contrasting in: (in)compatibility with manner adverbials (40); (no) event-related reading of quantifiers (cf. stative/degree interpretation in (41), in line with (29) above); (no) habitual reading in the present (42), as expected, and exactly as seen above in Romance.

- (40) *I kloun tromazun #(se) me sadhistiko/ aschimo tropo.*
 the clowns frighten 2s.ACC with sadistic rude manner
 ‘Clowns frighten you^{ok}(sadistically/rudely).’ (Transitive)
 ‘Clowns are scary #(sadistically/rudely).’ (SC)
- (41) *I sokolata pacheni ligho.*
 the chocolate fattens a bit
 ‘Chocolate is a little fattening.’
- (42) *I koka-kola #(se) fuskoni (kathe mera).*
 ‘Coke makes you bloat (every day).’ (Transitive)
 ‘Coke is bloat(en)ing’ (SC)

In Greek, perfective tenses in monadic frames force inchoative readings, like in Spanish. Notice the lifetime effect in the SC reading, typically seen in (Individual-Level) stative predications.

23. Hence solving the problem that motivates Alexiadou & Anagnostopoulou’s postulation of a different class of (putative) inchoatives with unexpected morphological/semantic/syntactic behavior.

24. Interestingly, constructions like (i) are not interpreted by default as NO (anaphoric object: *patates*), but as SCs (even if odd in this context).

- (i) *O furnos kei tis patates? Ne, oli i*
 the oven.NOM burn.3s the.FEM.ACC potatoes.ACC? Yes, all the
furni kene.
 oven.NOM.PL burn.3p
 ‘Does the oven burn the potatoes?’
 ‘Yes, all ovens cause burns’

- (43) *Afti i somba thermanthike.*
 ok^{ok}“This radiator heated up” (Inchoative)
 #[#]“This radiator had heating capacity” (SC)

Finally, SCs are productive with Object-Experiencer, but not with Subject-Experiencer verbs, in Greek as well.²⁵ Further patterns noted in both Greek and Romanian (Alexiadou & Iordăchioaia 2014: 63(15)–(19)) – *in/for-x-time*, spatial location, manner/instrumental/temporal modifiers – are correctly predicted by the present analysis, closely paralleling (27)–(33).

- (44) *Ta malja tu Kosta enoxlun.*
 the hair.PL the Kosta.GEN annoy.3P
 “Kosta’s hair is annoying.”

If the above treatment is on the right track, the problem – including syntactic, semantic and morphological distribution – consistently boils down to a derivational alternative. Therein lies its novelty. Apparently, however, languages may differ in the availability of this option, showing at least two patterns of variation:

- i. In languages where independent composition with v_{CAUSE} is systematically available, unique arguments of change-of-state verbs with zero alternation are ambiguous (Cause(r)/Undergoer readings available); otherwise (if available), special morphology generally marks composition with the unaccusative/dynamic V^o . In both cases, unmarked forms instantiate the Immediate Cause Linking Rule.
- ii. In a language where SC derivation is not systematic, the Default Linking Rule (Levin & Rappaport-Hovav 2000) holds (unique arguments are by default internal).²⁶ V^o defectiveness correlates with morphologically unmarked alternations.

25. Also here, subjects admitting only Cause(r) interpretation restrict the alternation. Otherwise, special morphology generally marks the realization of v_{BECOME} , with the consequent restriction on perfectivity and eventivity.

- (i) Object-Experiencer
O Mesi enolithike/#enohlise (grigora).
 the Messi.NOM annoyed.NAct/annoyed quickly
 “Messi got annoyed/#was annoying (quickly).”
 (cf. ^{ok}*O Mesi enoxli.* (the Messi annoys) “Messi is annoying.”)

Interestingly, Greek Subject-Experiencer verbs seem to lack stative forms and take special morphology (Alexiadou & Iordăchioaia 2014). This consistently extends the contrast with (i).

26. In this sense, it is not trivial to recall the observation about sole-argument-as-cause interpretation not being available in English ((6) above and footnotes).

In principle, English could instantiate pattern (ii), while Greek/Romance realize pattern (i). In this sense, an observation not to be overlooked is the contrast in the defective reading of unmarked monoargumental frames ((6) above requiring special morphology to equal the default (unaccusative) English reading (unmarked)).²⁷ This, we believe, is the crucial contribution of the present discussion.

3.3 Section summary

Section 3 strengthens the view that SCs are original monadic structures by extending the analysis to Romance psych verbs, a class of verbs which are known to have SC readings in other (e.g., Germanic) languages as well.

Greek data lends further support to the proposal. This includes the existence of stative variants of verbs participating in the alternation (psych verbs in Spanish, Marín & McNally 2011), along with the stative (or at least not defectively eventive) nature of the head introducing the external argument in the alternation (advanced by Ramchand 2008).

Moreover, Romance and Greek data strongly indicate that morphological defectiveness – the fact that unmarked (uncliticized) variants give SC readings – is consonant with Hale & Keyser's analysis of the causative monadic categorizer as a defective head in the alternation (see also Folli et al. 2005).²⁸ In this way, a transparent syntax/semantics/morphology correlation arises.

4. Final remarks

The idea of a minimal inventory of distinct external-argument-introducing *v* heads is an analytical choice not unanimously shared by constructionalist approaches. Yet, it has empirical advantages. Here I propose that the refinement – replacing Ramchand's *Init* for v_{CAUSE} in (35) – is necessary to accommodate several facts surrounding SCs.

A causative v^0 – in opposition to other external-argument introducing v^0 s like v_{DO} (Folli & Harley 2005), with each head restricting argument selection and interpretation (Folli & Harley 2007, among others) accordingly – straightforwardly

27. Cf. *Estos minerales corroen* 'These minerals are corrosive[\approx cause corrosion] vs *Estos minerales se corroen* 'These minerals corrode'[\approx become corroded]).

28. If correct, this would also be a case in which, where available, special morphology marks the transitivization of an otherwise intransitive verb (suggested by Alexiadou & Anagnostopoulou 2003, among others).

explains the patterns seen in SCs. Essentially, it would naturally accommodate the default interpretation of the subject as cause(r) independent of composition with the internal-argument-introducing V° , along with stative behavior. The solution follows Ramchand's (2008), among others, claim – in contrast to Folli & Harley – that in alternating verbs the projections surrounding $v_{\text{BECOME/Proc}}$ need not be eventive (see Mangialavori Rasia 2019 for extensive discussion). It also conforms to the observation that structures headed by the causative head give statives when not complemented by the internal-argument+event-introducing projection.

Verbs produced by v_{DO} select animate Agent subjects. Conversely, v_{CAUSE} – the head we are pointing to as the source of SCs – only requires that the *subject be a possible cause* (Folli & Harley 2005: 96). Such distinct selectional and interpretational restrictions would readily account for the contrast in interpretation of the subject in external-argument-only variants with otherwise similar configuration.²⁹ The distinction is relevant in the analysis as it could further predict the evident fact that not all unergative verbs select for animate (Agent) subjects, nor are properly represented by the active/agentive V_{DO} (notably, the whole class of unergative statives, cf. Levin & Rappaport-Hovav 2000).

Moreover, this opposition could explain the lack of productivity with verbs denoting a property of the external argument. With these verbs, transitive frames alternate with a variant involving an 'understood' object with generic interpretation (Levin 1993, among others) – vs SCs (original atransitives). Importantly, the verbs in question combine active denotation with agentive/intentional interpretation of the subject – clearly not interpretable as cause(r), as in SCs –, which is the expected pattern for v_{DO} .³⁰

(46) Characteristic Property of Agent Alternation.

a. *Ese perro muerde (gente).*

“That dog bites people.”

“That dog bites.”

SC

b. **Ese perro muerde.*

(intended) “That dog causes bite(s).”

29. Assuming that standard unergatives like *run* have a v_{DO} -headed structure like (35) (Hale & Keyser 2002; Folli & Harley 2007).

30. An important point here is that causative semantics is independent of agentive interpretation. As in other cases (see Rothmayr 2009, among others), the proposed differentiation is consistent with the claim that the external-argument-introducing head may receive causative interpretation without an agentive one. Significantly, only the occurrence of v_{DO} poses animacy restrictions like the ones setting a contrast between other (e.g., Property of Agent) alternations and SCs.

Eventivity and Agentivity tests show the expected distribution (recall (30)–(33) above) – note also the frequentative/habitual reading in (46a). Concomitantly, these properties mark a consequent semantic difference between SCs and transitivity alternations involving null/arb objects contrast. Here, the crosscut receives a structural explanation (VP-configuration).

- (47) a. *Lo que hace el perro es morder (voluntariamente/metódicamente).*
 “What the dog does is to bite (on purpose/methodically)”
- b. *#Lo que hace el payaso es asustar*
 (*#voluntariamente/#metódicamente*). (SC)
 (Intended) “What the clown does is to be frightening (voluntarily/methodically).”

Furthermore, the contrast between true (a)transitive alternations (SCs), and what eventually prove to be transitive constructions (null/arb Object alternations), also conforms to the generalization that verbs of surface contact (e.g., *bite*) and motion allow unspecified objects, while verbs of change of state do not (Levin & Rappaport-Hovav 2001, among others). The latter observation, widely embraced in the literature, lends further (lexical)semantic support³¹ to the analysis that SCs are original intransitive structures (as opposed to a null/arb object treatment).

To recap, Section 4 offers a contrast between SCs and structures associated with other external-argument heads proposed in the literature. Crucially, independent composition with v_{CAUSE} straightforwardly explains the default interpretation of the subject as cause(r) independent of the presence of the process V licensing an internal argument in SCs, at the same time that it captures an asymmetry between two thus far undifferentiated (a)transitivity alternations. Hence, a contrast arises between apparent intransitive forms with null objects (pro/Arb-Object alternation) and true monadic intransitives (SCs).³² Moreover, the independent status of this head in the composition could readily capture its anticipated productivity in stative constructions in absence of the process/internal-argument-introducing head.

31. Insofar as (lexical-semantic) verb meaning is correctly considered to influence syntactic behavior.

32. A final answer to the specific parameters that dissociate between the external-argument-introducing verbs is left for further work.

5. Conclusions

Several facts suggest that in languages like Romance and Greek, the derivation of eventless, external-argument-only monadic causatives is a free/systematic derivational option. Following previous accounts of atransitive (McIntyre 2004) and Stative-Causative (Ramchand 2008) constructions, I propose that this is done via direct composition with the causative external-argument-introducing v^o (v_{CAUSE}), under the condition that v^o is not complemented by the internal-argument-introducing head ($V^o_{\text{Proc/BECOME}}$). The result is a causative syntactically, morphologically and semantically simpler than the (dyadic) causative commonly analyzed, and maximally distinct from the better-known (inchoative) monadic variant. Apparently, languages may differ in the availability of this option, crucially revealing two distinct patterns of variation.

Acknowledgements

Thanks go to Giorgios Spathas and M. Mavrogiorgos for Greek data. Also, to colleagues, audiences as well as the anonymous reviewers for their insightful questions. Any errors are my own.

References

- Alexiadou, Artemis & Elena Anagnostopoulou. 2003. Voice morphology in the causative–inchoative alternation. In Artemis Alexiadou, Elena Anagnostopoulou & Martin Everaert (eds.), *The unaccusativity puzzle*, 115–136. Oxford: Oxford University Press.
- Alexiadou, Artemis, Elena Anagnostopoulou & Florian Schäfer. 2006. The properties of anti-causatives crosslinguistically. In Mara Frascarelli (ed.), *Phases of interpretation*, 187–211. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110197723.4.187>
- Alexiadou, Artemis & Gianina Iordăchioaia. 2014. The psych causative alternation. *Lingua* 148. 53–79. <https://doi.org/10.1016/j.lingua.2014.05.010>
- Arad, Maya. 1998. *VP-Structure and the syntax-lexicon interface*. Cambridge, MA: MIT Press.
- Bhatt, Rajesh & Roumyana Pancheva. 2006. Implicit arguments. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, Volume 2, 558–588. Malden, MA: Blackwell. <https://doi.org/10.1002/9780470996591.ch34>
- Cattaneo, Andrea. 2008. Italian null objects and resultative/depictive predication. In Emily Elfner & Martin Wakov (eds.), *Proceedings of NELS 37*, 141–152. Amherst: GLSA Publications.
- Chierchia, Gennaro. 2004. A semantics for unaccusatives and its syntactic consequences. In Artemis Alexiadou, Elena Anagnostopoulou & Martin Everaert (eds.), *The unaccusativity puzzle*, 22–59. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199257652.003.0002>
- Doron, Edit. 2003. Agency and voice: The semantics of the Semitic templates. *Natural Language Semantics* 11. 1–67. <https://doi.org/10.1023/A:1023021423453>

- Dowty, David. 1979. *Word meaning and Montague grammar*. Dordrecht: Reidel.
<https://doi.org/10.1007/978-94-009-9473-7>
- Folli, Raffaella & Heidi Harley. 2005. Flavors of v. In Roumyana Slabakova & Paula Kempchinsky (eds.), *Aspectual inquiries*, 95–120. Dordrecht: Kluwer.
https://doi.org/10.1007/1-4020-3033-9_5
- Folli, Raffaella & Heidi Harley. 2007. Causation, obligation and argument structure. *Linguistic Inquiry* 38. 207–249. <https://doi.org/10.1162/ling.2007.38.2.197>
- Folli, Raffaella, Heidi Harley & Simin Karimi. 2005. Determinants of event type in Persian complex predicates. *Lingua* 115. 1365–1401. <https://doi.org/10.1016/j.lingua.2004.06.002>
- Hale, Ken & Samuel J. Keyser. 1988. Explaining and constraining the English middle. In Carol Tenny (ed.), *Studies in Generative Approaches to Aspect*, 41–58. Cambridge, MA: MIT Press.
- Hale, Ken & Samuel J. Keyser. 1993. On the argument structure and the lexical expression of syntactic relations. In Ken Hale & Samuel Keyser (eds.), *The view from Building 20*, 53–109. Cambridge, MA: MIT Press.
- Hale, Kenneth; Samuel Jay Keyser. 1998. The Basic Elements of Argument Structure. *MIT Working Papers in Linguistics* 32, 73–118.
- Hale, Ken & Samuel J. Keyser. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/5634.001.0001>
- Harley, Heidi. 2012. Lexical decomposition in modern syntactic theory. In Wolfram Hinzen, Edouard Machery & Markus Werning (eds.), *The Oxford handbook of compositionality*, 328–350. Oxford: Oxford University Press.
- Harley, Heidi. 2014. On the identity of roots. *Theoretical Linguistics* 40. 225–276.
<https://doi.org/10.1515/tl-2014-0010>
- Harley, Heidi. 2017. The “bundling” hypothesis and the disparate functions of little v. In Roberta D’Alessandro, Irene Franco & Ángel J. Gallego (eds.), *The verbal domain*, 3–28. Oxford: Oxford University Press.
- Haspelmath, Martin. 1993. More on the typology of inchoative/causative verb alternations. In Bernard Comrie & Maria Polinsky (eds.), *Causatives and transitivity*, 87–120. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.23.05has>
- Jackendoff, Ray. 1990. *Semantic structures*. Cambridge, MA: MIT Press.
- Jackendoff, Ray. 1991. Parts and boundaries. *Cognition* 41. 9–45.
[https://doi.org/10.1016/0010-0277\(91\)90031-X](https://doi.org/10.1016/0010-0277(91)90031-X)
- Levin, Beth. 1993. *Verb classes in English*. Chicago: University of Chicago Press.
- Levin, Beth & Malka Rappaport-Hovav. 1995. *Unaccusativity*. Cambridge, MA: MIT Press.
- Levin, Beth & Malka Rappaport-Hovav. 2000. Classifying single argument verbs. In Peter Coopmans, Martin Everaert & Jane Grimshaw (eds.), *Lexical specification and insertion*, 269–304. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.197.13hov>
- Levin, Beth & Malka Rappaport-Hovav. 2001. An event structure account of English resultatives. *Language* 77. 766–797. <https://doi.org/10.1353/lan.2001.0221>
- Levin, Beth & Malka Rappaport-Hovav. 2010. Lexicalized scales and verbs of scalar change. Paper presented at the 46th Annual Meeting of the CLS, Chicago, IL, April 8–10.
- Levin, Beth & Malka Rappaport-Hovav. 2012. Lexicon uniformity and the causative alternation. Martin Everaert, Marijana Marelj & Tal Siloni (eds.), *The theta system*, 1–19. Oxford: Oxford University Press.
- Levin, Beth. Manuscript. *English object alternations: A unified account*. Stanford University. 2006.
- Maienborn, Claudia. 2007. On Davidsonian and Kimian states. In Ileana Comorovski & Klaus von Heusinger (eds.), *Existence: Syntax and semantics*, 107–130. Dordrecht: Kluwer.

- Mangialavori Rasia, Maria Eugenia. 2019. Stativity in the causative alternation? New questions and a new variant. *Open Linguistics* 5. 233–259. <https://doi.org/10.1515/opli-2019-0014>
- Mangialavori & Marin. 2021. Psych verb transitivity: the English/Romance crosscut. Talk at Sensus 2. Constructing meaning in Romance. October, 1, 2021, UMass, Amherst.
- Marín, Rafael & Louise McNally. 2011. Inchoativity, change of state, and telicity: Evidence from Spanish reflexive psychological verbs. *Natural Language and Linguistic Theory* 29. 467–502. <https://doi.org/10.1007/s11049-011-9127-3>
- Massam, Diane. 1989. Null objects and non-thematic subjects in middles and tough constructions. *Toronto Working Papers in Linguistics* 10. 95–118.
- Massam, Diane. 1992. Null objects and non-thematic Subjects. *Journal of Linguistics* 28. 115–137. <https://doi.org/10.1017/S0022226700015012>
- McIntyre, Andrew. 2004. Event paths, conflation, argument structure and VP shells. *Linguistics* 42. 523–571. <https://doi.org/10.1515/ling.2004.018>
- Piñón, Christopher. 2001. A finer look at the causative-inchoative alternation. In Rachel Hastings, Brendan Jackson & Zsófia Zvolenszky (eds.), *Proceedings of Semantics and Linguistic Theory XI*, 346–364. Ithaca: CLC Publications. <https://doi.org/10.3765/salt.v11i0.2858>
- Ramchand, Gillian. 2008. *Verb meaning and the lexicon: A first-phase syntax*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486319>
- Reinhart, Tanya. 2002. The theta system: An overview. *Theoretical Linguistics* 28. 229–290.
- Rizzi, Luigi. 2003. *Comparative syntax and language acquisition*. London: Routledge. <https://doi.org/10.4324/9780203461785>
- Roberts, Ian. 1987. *The representation of implicit and dethematized subjects*. Dordrecht: Foris. <https://doi.org/10.1515/9783110877304>
- Rosen, Susan. 1996. Events and verb classification. *Linguistics* 34. 191–223. <https://doi.org/10.1515/ling.1996.34.2.191>
- Rothmayr, Antonia. 2009. *The structure of stative verbs*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.143>
- Rothstein, Susan. 2012. Another look at accomplishments and incrementality. In Violeta Demonte & Louise McNally (eds.), *Telicity, change and state*, 60–102. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199693498.003.0003>
- Russi, Cinzia. 2008. *Italian clitics*. Berlin: de Gruyter. <https://doi.org/10.1515/9783110206975>
- Zubizarreta, Maria Luisa & Eunjeong Oh. 2007. *On the syntactic composition of manner and motion*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/5132.001.0001>

On *wh*-extraction in *de+que* constructions in Spanish

Gabriel Martínez Vera
Goethe-Universität Frankfurt

This chapter discusses *wh*-extraction in Spanish clauses involving a *V+de+CP* sequence. The previous literature observes that objects, not adjuncts, may be extracted from a *de+CP* clause. I make three novel observations: (i) subjects pattern with adjuncts in that they cannot be extracted from *de+CP*, (ii) there is a ‘distance’ effect in that subject extraction improves with one more level of embedding, and (iii) there is also a ‘distance’ effect regarding adjunct extraction with one more level of embedding. (i)–(iii) are surprising, since subjects otherwise do not pattern with adjuncts regarding extraction in Spanish, and adjunct extraction out of islands is otherwise unacceptable. I provide a phasal account where Phase Collapsing and successive cyclic movement are crucial.

Keywords: *wh*-extraction, islands, phases, incorporation, Spanish

1. Introduction

There is a well-known object vs adjunct contrast regarding *wh*-extraction from islands in English. Although both are disallowed, adjunct extraction is worse than object extraction (Huang 1982; Chomsky 1986; among others), as (1) illustrates with *whether*-islands (throughout the chapter, I focus on the relative grammaticality difference between examples).

- (1) a. [?]**What did you wonder [whether John fixed t]?* *object*
b. **How did you wonder [whether John fixed the car t]?* *adjunct*

Subject extraction is as ungrammatical as adjunct extraction:

- (2) **Who did you wonder [whether t fixed the car]?* *subject*

However, subject extraction improves with ‘distance’, i.e., when extraction takes place from a more embedded position: it patterns with object extraction – object and adjunct extraction remain the same.

- (3) a. [?]**What did you wonder [whether Mary said [John fixed t]]?* *object*
 b. [?]**Who did you wonder [whether Mary said [t fixed the car]]* *subject*
 c. **How did you wonder [whether Mary said [John fixed the car t]]?* *adjunct*

The reason why subject extraction improves in (3b) when compared to (2) lies in whether extraction takes places from right below the island phrase. It takes place from Spec, IP right below the island CP in (2) – but not in (3b). That is why subjects pattern with objects in (3b) but with adjuncts in (2).

In Spanish, object and adjunct extraction from islands is also ungrammatical (Uriagereka 1988; Gallego 2007; Saab manuscript). As in English, object extraction is less degraded than adjunct extraction. I use *because*-islands to illustrate.

- (4) a. [?]¿*Qué te reíste [porque arregló Juan t]?* *object*
 what CL laughed because fixed John
 “What did you laugh because John fixed?”
 b. [?]¿*Cómo te reíste [porque arregló Juan el auto t]?* *adjunct*
 how CL laughed because fixed John the car
 “How did you laugh because John fixed the car?”

Subjects, unlike English (1a)–(2), pattern with objects. This is because Spanish allows subject extraction from Spec, *v*P.

- (5) [?]¿*Quién te reíste [porque arregló t el auto]?* *subject*
 who CL laughed because fixed the car
 “Who did you laugh because fixed the car?”

There is no ‘distance’ effect in Spanish (6): subjects pattern with objects (i.e., (5) already instantiates the distance effect, since subjects never move from right below the island CP); subject extraction is less degraded than adjunct extraction here as well.

- (6) a. [?]¿*Qué te reíste [porque dijo María [que arregló Juan t]]?*
 what CL laughed because said Mary that fixed John
 “What did you laugh because Mary said John fixed?” *object*
 b. [?]¿*Quién te reíste [porque dijo María [que arregló t el auto]?*
 what CL laughed because said Mary that fixed the car
 “Who did you laugh because Mary said fixed the car?” *subject*
 c. [?]¿*Cómo te reíste [porque dijo María [que arregló Juan el auto t]?*
 how CL laughed because said Mary that fixed John the car
 “How did you laugh because Mary said John fixed the car?” *adjunct*

Thus, in English, subjects pattern with adjuncts in the ‘no-distance’ case and with objects in the ‘distance’ case. In Spanish, subjects pattern with objects in the ‘no-distance’ *and* ‘distance’ cases; adjuncts pattern differently.

It is also relevant to introduce Spanish embedded clauses. As in English, verbs take CPs as complements (7). This V+CP sequence is a general strategy in the language.

- (7) *María cree [que Juan arregló el auto].*
 Mary thinks that John fixed the car
 “Mary thinks that John fixed the car.”

Spanish also has verbs that take embedded clauses with the preposition *de* “of” preceding the CP, i.e., a V+*de*+CP sequence (8) (Demonte & Fernández Soriano 2005; Bošković 2015).¹

- (8) *Te convenciste [de que Juan arregló el auto].*
 CL convinced of that Juan fixed the car
 “You convinced yourself that John fixed the car.”

This chapter focuses on *wh*-extraction out of V+*de*+CP. These clauses bring in islandhood (Bošković 2015): like English (1) and Spanish (4), (9) shows that adjunct extraction is ungrammatical.

- (9) *¿Cómo te convenciste [*de que arregló Juan el auto* t]? *adjunct*
 how CL convinced of that fixed John the car
 “How did you convince yourself that John fixed the car?”

However, extraction from V+*de*+CP is not disallowed across the board: Demonte & Fernández Soriano (2005: 1078) show that object extraction is possible:²

1. I only focus on finite CPs. I use *convencer* “convince” in the examples; other verbs are *acusar* “accuse”, *dudar* “doubt”, *hablar* “talk”, etc.

2. Bošković (2015) reports that object extraction in cases like (10) is degraded. Here I provide an account of the judgment reported in Demonte & Fernández Soriano (2005). In my data survey, I found that judgments are split among speakers: some share the judgment Bošković reports; some share the judgment Demonte & Fernández Soriano report. In particular, I have gathered judgments from different consultants, which show that there is indeed a split among speakers, with some allowing and some disallowing object extraction in cases like (10) (confirming the split between the judgment reported by Bošković 2015 and Demonte & Fernández Soriano 2005). For contexts where extraction is allowed, as in cases like (10), I collected the judgments from 8 consultants (2 from Spain, 2 from Colombia, 3 from Peru and 1 from Argentina). All stimuli were introduced orally controlling for possible prosodic confounding factors. There were also speakers I consulted who did not share the contrasts. Specifically, there were 4 consultants who did not allow *wh*-extraction in the cases under discussion (1 from Spain, 1 from Ecuador, 1 from Peru, and 1 from Argentina). The methodology noted above was used with these consultants as well. While the data collection follows common practice in theoretical linguistics, given that the sample is small (12 speakers overall) and pending a systematic study from an experimental perspective, this difference in proportion (2 to 1) should not be understood in terms of there being more speakers

- (10) *¿Qué te convenciste [de que arregló Juan t]?* object
 what CL convinced of that fixed John
 “What did you convince yourself that John fixed?”

Interestingly, subject extraction, which has not been discussed before, is ungrammatical.³

- (11) **¿Quién te convenciste [de que arregló t el carro]?* subject
 who CL convinced of that fixed the car
 “Who did you convince yourself that fixed the car?”

Subject extraction (11) is as ungrammatical as adjunct extraction (9). This is like English (1b)–(2) and unlike Spanish (4b)–(5).

Actually, Spanish V+*de*+CP is similar to English in another respect, i.e., the ‘distance’ case. With one more level of embedding, subject extraction improves, likening it to object extraction.

- (12) a. *¿Qué te convenciste [de que dijo María [que arregló Juan t]?* object
 what CL convinced of that said Mary that fixed John
 “What did you convince yourself that Mary said that John fixed?”
 b. *¿Quién te convenciste [de que dijo María que arregló t*
 who CL convinced of that said Mary that fixed
el auto t]? subject
 the car
 “Who did you convince yourself that Mary said that fixed the car?”

There is one more novel observation I make. In V+*de*+CP, there is a ‘distance’ effect regarding adjunct extraction: it improves with one more level of embedding (13) – it improves as much as subject extraction (12b). This is surprising, as such an effect has not been attested in the literature (see English (3c) and Spanish (6c)).

- (13) *¿Cómo te convenciste [de que dijo María [que arregló Juan el auto? adjunct*
 how CL convinced of that said Mary that fixed John the car
 “How did you convince yourself that Mary said that John fixed the car?”

who allow *wh*-extraction. What this shows is merely that there clearly is a split among speakers. This is what the reader should bear in mind; note also that the split does not seem to correlate precisely with any geographical divisions. It should also be noted that the phenomena discussed here are not instances of *dequeísmo*, a separate phenomenon analyzed by Demonte & Fernández Soriano (2005). *Dequeísmo* refers to a phenomenon that is present in some dialects of Spanish (in both Peninsular and Latin American dialects) involving hypercorrection in that, in cases where speakers of the standard dialect(s) would use a CP (headed by the C *que*), speakers of these other dialects make use of *de* + CP (in this regard, note the presence of *de* + *que* in *de-que-ísmo*).

3. The asymmetries in (9)–(11) do not change with subjunctive CPs.

This chapter accounts for these observations, summarized below: (i) subjects pattern with adjuncts in that they cannot be extracted from a single *de*+CP clause, (ii) there is a ‘distance’ effect in that subject extraction improves when one more level of embedding is added, and (iii) there is also a ‘distance’ effect regarding adjunct extraction when one more level of embedding is added.

Section 2 accounts for observations (i)–(iii). The proposal adopts the phasal framework. I make use of the mechanism of Phase Collapsing (Bošković 2015, 2016), whereby two phase heads/phases are collapsed into one. Successive cyclic movement is also crucial. Based on Nunes’s (2014, 2016), I propose a mechanism that makes subjects and adjuncts pattern together in terms of movement, which accounts for why extraction is disallowed with a single *de*+CP clause, but leaves room for improvement in the ‘distance’ case.

2. Proposal

I propose a phasal account of the phenomena under discussion. Section 2.1 introduces Phase-over-Phase Configurations and Phase Collapsing. Section 2.2 presents Nunes’s (2014, 2016) approach to successive cyclic movement. Section 2.3 derives the extraction patterns from Section 1.

2.1 Phase-over-phase configurations and phase collapsing

Chomsky (2000, 2001) proposes a rigid approach to phasehood where CPs and *v*Ps are always phases. More recently, other approaches define phasehood contextually: the phasal status of an element depends on the syntactic context where this element is (Bošković 2005, 2014; den Dikken 2007; Gallego & Uriagereka 2007). I assume one such approach, i.e., Bošković (2015, 2016): there are two structural domains, the thematic and the non-thematic, and that the highest projection in each domain is a phase. This often corresponds to Chomsky’s original approach: *v*P is a phase, since it is the highest projection in the (verbal) thematic domain; CP is a phase, since it is the highest projection in the non-thematic domain.

Of particular interest are Phase-over-Phase Configurations, which, as Bošković (2015, 2016) discusses, are structures where two phasal projections are stacked one on top of the other (14).⁴

4. Due to space constraints, the account is limited to extraction from one particular island configuration, where two phases are stacked one on top of the other. I do not discuss other kinds of islands, where other issues disallowing extraction may be involved, e.g., in (1)–(6).

(14) *Phase-over-Phase Configuration*

$$[_{YP=Phase} [_{XP=Phase}]]$$

Bošković further shows that extraction out of the lower phase in (14) is disallowed (15).

(15) $[\dots [_{YP=Phase} [_{XP=Phase} \dots ZP \dots]]]$

Bošković calls the ban on extraction from the phasal complement of a phase head the Phase-over-Phase Constraint.

(16) *Phase-over-Phase Constraint*

Extraction is banned from phases that function as complements of phase heads.

Bošković's initial motivation for (16) is the Complex NP Constraint, which bans the extraction from the clausal complement of N. (17) involves extraction out of a CP, the complement of the N *rumors*. NP and C are the highest projections in their phasal domains, so they are phases. Thus, this is a Phase-over-Phase Configuration, which is subject to the Phase-over-Phase Constraint. Extraction out of the CP is banned – in line with this, extracting *how* from the CP in (17) is not possible.

(17) **How did you hear* [_{NP=Phase} *rumors* [_{CP=Phase} *that Sue bought a house t*]]?

Bošković argues that the Complex NP Constraint holds more generally for all complements of nouns, and extends it to cases involving adjectives and prepositions (this is the Complex XP Constraint). He shows that extraction out of the complement of As and Ps is disallowed (see Bošković 2015 for additional cases). (18) shows this: *how* (18a) and *of whom* (18b) cannot be extracted from phases that are themselves complements of phases.

(18) a. ?**How is he* [_{AP=Phase} *proud* [_{CP=Phase} *that Bill hired John t*]]?
 b. ?**Of who(m) did you read* [_{PP=Phase} *about* [_{DP=Phase} *friends t*]]?

Bošković notes, however, that Phase-over-Phase Configurations do not always disallow extraction. He argues that the Phase-over-Phase Constraint can be circumvented under Phase Collapsing (19), which is a reanalysis process where two phasal heads/phases are reanalyzed together. The lower phase head incorporates into the higher one, which voids the phasehood of the lower phase.

(19) *Phase Collapsing*

$$[_{YP=Phase} [_{XP=Phase}]] \rightarrow [_{YP=Phase} Y+X_i [_{XP} t_i]]$$

Bošković exemplifies this mechanism with a number of languages, and, within Romance, with Galician. Galician has D-incorporation which voids islandhood

(Uriagereka 1988; Bošković 2013). Galician disallows movement from definite DPs, but this violation can be voided when D incorporates into the verb (following Bošković 2015, I assume that V moves to ν and D incorporates into $\nu+V$, where traces do not count as interveners). In (20), the object of the N is extracted.

- (20) a. **e de quén viche* [DP *o retrato* t]?
 and of who saw(you) the portrait
 b. **e de quén_j viche-lo_i* [DP *t_i retrato t_j*?]
 and of whom say(you)-the portrait
 “so, who have you seen the portrait of?” (Uriagereka 1988: 81)

Bošković assumes that (some) traditional islands, including definite DPs, disallow movement through their edge (which is the source of their islandhood). ν cannot attract the *wh*-object in (20a) due to the Phase Impenetrability Condition (PIC), which prevents extraction from the complement of phase XP (the DP) (see also Section 2.2). In (20b), there is a complex phase due to Phase Collapsing (i.e., D-incorporation), so DP is no longer a phase. The *wh*-object can thus move directly to the edge of the ν P phase.

Bošković further shows that, in contrast to object extraction, adjunct extraction remains disallowed under Phase Collapsing. In Galician, while object extraction is possible under Phase Collapsing (20b), adjunct extraction is not, regardless of whether Phase Collapsing takes place (21).⁵

- (21) a. **Por quem escoitamos* [DP *a descripcion* t]?
 by whom listened(we) the description
 b. **Por quem_j escoitamo-la_i* [DP *t_i descripcion t_j*?]
 by whom listened(we)-the description
 “By whom did we listen the description” (Bošković 2015: 643)

Although Bošković notes the contrast between object and adjunct extraction under Phase Collapsing, he leaves it as a puzzle. Section 2.3 provides an account of $V+de+CP$ that derives this contrast.

2.2 Successive cyclic movement

I also adopt Nunes’s (2014, 2016) approach to successive cyclic movement, which builds on Chomsky (2000, 2001) and Bošković (2007). The latter are briefly discussed before turning to how Nunes incorporates them into his proposal.

Chomsky (2000, 2001) proposes that long distance movement proceeds successive cyclically in consonance with the PIC (only the head and the Spec of a phase

5. Subject extraction could not be tested in the cases Bošković (2015) discussed.

are accessible for movement out of a phase). He instantiates this movement by (optionally) assigning an EPP feature to phase heads, which is satisfied by movement to the phase edge, so the moved element is accessible for movement out of the phase. Focusing on the embedded C in (22), *that* has an EPP feature; *what* moves to the edge of the phase on its way to the matrix CP. If *that* did not have an EPP feature, *what* could not move to the embedded CP edge, hence could not move out of this CP due to the PIC.

(22) *What did Mary think [that John bought]?*

This approach, however, faces problems with cases like (23). Focusing on the embedded C, nothing bans *that* from having an EPP feature, with *what* moving to the edge of this phase.

(23) **Who thinks what that Mary bought?*

To address this issue, Chomsky suggests that *that* can be assigned an EPP feature only when necessary to allow successive cyclic movement, hence not in (23). As Bošković (2007) points out, this faces a look-ahead problem when considering (22)–(23) derivationally. He focuses on the step of the derivation in (24). At this point, (22)–(23) are identical; it will only be known whether *what* needs to move successive cyclically when the structure is further expanded. In (24), in order to assess whether *that* can be assigned an EPP feature we need to know how the structure above this CP will be expanded, i.e., whether further movement of *what* will be needed. It is in (22), but it is not in (23). This is the look-ahead problem Bošković notes.

(24) [_{CP} *what* [*that John bought t*]]

Bošković (2007) takes a different approach to successive cyclic movement. Instead of linking successive cyclic movement to phase heads, he links it to the moving elements. He proposes that successive cyclic movement is triggered by an uninterpretable feature *uF* in the moving element (which happens to be optional with *wh*-phrases in English). This feature needs checking by a relevant interpretable feature *iF* (here, it is an interrogative feature). Bošković further argues that *uF* in general must be a probe; it must c-command the relevant *iF*. The element with *uF* keeps moving to phase edges to avoid being sent to spell-out until it finds itself in a position where it c-commands *iF*. *What* then has to move out of Spec,*vP*, since its *uF* feature, which triggers movement to the phasal edge, cannot be checked there.

There is no look-ahead problem in this case: there is no need to know how the structure in (24) will be expanded. So, in (25a), *what* must have had the relevant *uF* feature or it could not move to Spec,CP. In (25b), the *uF* of *what* cannot be checked

(it does not move to the matrix CP, as *uF* must c-command the corresponding *iF*), so the derivation crashes.

- (25) a. **What_i did Mary think [t_i that John bought]_i?*
 b. **Who thinks what that Mary bought?*

Nunes (2014, 2016) points out, however, that phase heads may play a role in allowing/disallowing successive cyclic movement. One case for him is the English *Comp*-trace effect, where the presence of an overt vs covert C is relevant. Thus, in (26), the presence/absence of *that* correlates with the impossibility/possibility of subject extraction.

- (26) *Who do you think (*that) saw Bill?*

Nunes's solution is to incorporate Chomsky's intuition that *wh*-movement may be linked to phase heads in addition to adopting Bošković's system. He proposes a hybrid system where *uF* is associated with the moving element, like in Bošković's system, but the source of *uF* could be lexically associated with the moving element from the beginning (Bošković 2007) or can be lexically associated with phase heads (Chomsky 2001). If *uF* is associated with phase heads, *uF* will be assigned to the element that will be moving. His proposal is as stated below:

- (27) An uninterpretable feature triggering movement *uF* can be lexically encoded on:
 a. *wh*-elements; or
 b. phase heads.

If the latter is the case, the phase head may assign *uF* to an element in its probe domain (i.e., to an element within the complement of the phase head).

As Nunes adopts Bošković's system, this account does not face the look-ahead problem either. (27b) is the relevant case here – English is a language in which (27b) holds.⁶ It is schematically represented in (28) – X is a phase head and the *wh*-element is in its probe domain. Here, the *wh*-element can move.

- (28) [_{XP} ... X_[*uF*] ... *wh*-element] → [_{XP} ... X ... *wh*-element_[*uF*]]

Nunes's account of (22)/(25a)–(23)/(25b) is like Bošković's, the difference being that the moving element, *what*, gets *uF* from a phase head, i.e., *v* here.

Recall (26), where the presence/absence of an overt C correlates with the impossibility/possibility of extracting the subject. Nunes's system makes this explicit.

6. See Nunes (2014, 2016) for extensive discussion regarding (27a–b), which is part of Nunes's proposal to account for cross-linguistic variation.

He argues that the possibility of assigning uF is tied to the different Cs in (26): *that* is not specified for uF , whereas its null counterpart C_{null} is optionally specified for uF . Thus, C_{null} but not *that* can assign uF to the subject in this case. This accounts for the correlations in (26): with *that*, subject extraction is not possible; with C_{null} , it is.

2.3 Deriving the asymmetries in *wh*-extraction patterns in $V+de+CP$

The discussion about Phase-over-Phase Configurations and Phase Collapsing in Section 2.1 constitutes the baseline to derive the asymmetries in *wh*-extraction patterns in $V+de+CP$. (29) repeats the ‘no-distance’ Examples (9)–(11):

- (29) a. *¿Qué te convenciste [*de que* arregló Juan t]? object
 what CL convinced of that fixed John
 “What did you convince yourself that John fixed?”
 b. *¿Quién te convenciste [*de que* arregló t el carro]? subject
 who CL convinced of that fixed the car
 “Who did you convince yourself that fixed the car?”
 c. *¿Cómo te convenciste [*de que* arregló Juan el auto t]? adjunct
 how CL convinced of that fixed John the car
 “How did you convinced yourself that John fixed the car?”

I assume the structure in (30) for $de+CP$ – assuming that *de* is a P (Demonte & Fernández Soriano 2005). (30) is a Phase-over-Phase Configuration.

- (30) [_{PP=Phase} *de* [_{CP=Phase} *que*]]

Leaving the structure like this would be too strong, since extraction out of $de+CP$ is not banned across the board.

I pursue an analysis where *de* and *que* are collapsed, i.e., *que* is not a phase head anymore. This means that *que* moves to *de*.

- (31) [_{PP+CP=Phase} *de+que*]



Recall that Bošković (2015, 2016) shows that Phase Collapsing improves object but not adjunct extraction, which is the case in the ‘no-distance’ examples in $V+de+CP$. I have added that subjects pattern with adjuncts in $V+de+CP$ to this.

I derive these extraction asymmetries in $V+de+CP$ by means of an implementation of Phase Collapsing combined with Nunes’s (2014; 2016) approach to successive cyclic movement, specifically, case (27b), where phase heads assign uF to the element that will move (this takes place in the probe domain of the relevant phase head). Recall that, in Phase Collapsing, there is head movement of the lower head (*que* here) to the higher head (*de* here). I propose that movement of *que* is triggered by a feature, deF , which is also present in *de*. deF indicates locally that *que*


will move to *de*, which means that its phasal status will be voided, hence spell-out of the complement of *que* does not take place. PP is still a phase. Moreover, I assume *deF* is incompatible with assigning *uF* (the feature triggering movement), i.e., whenever *deF* is present, *uF* cannot be.

With this setup, I account for the observations (i)–(iii) in Section 1: (i) subjects pattern with adjuncts in that they cannot be extracted from a *de+CP* clause, (ii) there is a ‘distance’ effect in that subject extraction improves when one more level of embedding is added, and (iii) there is also a ‘distance’ effect regarding adjunct extraction when one more level of embedding is added.

The account of (i) is as follows. Subjects and adjuncts are in the probe domain of the *de+que* phase head. This complex phase head cannot assign *uF* (which is incompatible with *deF*). Thus, subjects and adjuncts cannot move. This is shown in (32). The matrix C cannot check its interrogative feature, so the derivation crashes. Subjects and adjuncts thus pattern together: their extraction is disallowed. This is the case of (29b–c).

- (32) a. [... [_{PP+CP=Phase} *de+que* ... [_{VP} *wh*-subject *v*...]]]

 b. [... [_{PP+CP=Phase} *de+que* ... [... *v* ...] *wh*-adjunct]]


Objects, however, remain unaffected. The phase head with *uF* is *v*, which assigns it to the object – as it is in its probe domain (since subjects and adjuncts are base-generated above *v*, they are not in the probe domain of *v*). Thus, the object can move, leaving the *de+que* phase on its way to the matrix C to check *uF*. This is the case of (29a).


- (33) a. [... [_{PP+CP=Phase} *de+que* ... [_{VP} ... *v*_[uF] *wh*-object]]] →
 [... [_{PP+CP=Phase} *de+que* ... [_{VP} ... *v* *wh*-object_[uF]]]]
 b. [... [_{PP+CP=Phase} *de+que* ... [_{VP} ... *v* *wh*-object_[uF]]]]


I now turn to (ii)–(iii) (the ‘distance’ cases). (34) repeats (12)–(13):

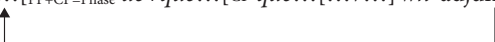
- (34) a. ¿Qué te convenciste [*de que* dijo María [*que* arregló Juan t]? *object*
 what CL convinced of that said Mary that fixed John
 “What did you convince yourself that Mary said that John fixed?”
 b. ¿Quién te convenciste [*de que* dijo María *que*
 who CL convinced of that said Mary that
arregló t el auto t]? *subject*
 fixed the car
 “Who did you convince yourself that Mary said that fixed the car?”

- c. ¿Cómo te convenciste [*de que* dijo María [*que*
 how CL convinced of that said Mary that
arregló Juan el auto? *adjunct*
 fixed John the car
 “How did you convince yourself that Mary said that John fixed the car?”


The account of (ii), the case where subject extraction improves with ‘distance’, is as follows. The lower embedded clause is headed by the C *que* (this head does not have *de*F, as it does not incorporate into *de*). This head can in fact assign *u*F in its probe domain, specifically, it can assign *u*F to the subject. Thus, the subject can move; it can then leave the *de+que* phase and move all the way up to the matrix C. This is shown in (35). This accounts for the improvement of subject extraction in the ‘distance’ case (34b).

- (35) a. [... [PP+CP=Phase *de+que* ... [CP *que*_[uF] ... [vP *wh*-subject *v* ...]]] →
 [... [PP+CP=Phase *de+que* ... [CP *que* ... [vP *wh*-subject_[uF] *v* ...]]]
 b. [... [PP+CP=Phase *de+que*...[CP *que*...[vP *wh*-subject_[uF]*v*...]]]


The account of (iii), the case where adjunct extraction improves with ‘distance’, is like (35). Lower *que* can assign *u*F in its probe domain, the adjunct here. The adjunct can then move, eventually reaching the matrix C to check *u*F. This is shown in (36). This accounts for the ‘distance’ improvement of adjunct extraction (34c).

- (36) a. [... [PP+CP=Phase *de+que* ... [CP *que*_[uF] ... [... *v* ...] *wh*-adjunct]]] →
 [... [PP+CP=Phase *de+que* ... [CP *que* ... [... *v* ...] *wh*-adjunct_[uF]]]]
 b. [... [PP+CP=Phase *de+que*...[CP *que*...[...*v*...] *wh*-adjunct_[uF]]]]


As expected, object extraction with ‘distance’ (34c) is also possible. *v* in the lower CP can assign *u*F to the object, which can reach the matrix C.

- (37) a. [... [PP+CP=Phase *de+que* ... [CP ... [vP ... *v*_[uF] *wh*-object]]]] →
 [... [PP+CP=Phase *de+que* ... [CP ... [vP ... *v* *wh*-object_[uF]]]]]]
 b. [... [PP+CP=Phase *de+que* ... [CP ... [vP ... *v* *wh*-object_[uF]]]]]]


The analysis derives the asymmetries under Phase Collapsing noted by Bošković (2015, 2016). It accounts for why subjects pattern with adjuncts in *V+de+CP*, while leaving room for improvement under ‘distance’.

3. Conclusion

I discussed the asymmetries in *wh*-extraction in Spanish clauses involving V+*de*+CP. I made three novel observations: (i) subjects pattern with adjuncts in that they cannot be extracted from a *de*+CP clause, (ii) there is a ‘distance’ effect in that subject extraction improves when one more level of embedding is added, and (iii) there is also a ‘distance’ effect regarding adjunct extraction with one more level of embedding. (i)–(iii) are surprising, since subjects otherwise do not pattern with adjuncts regarding extraction in Spanish (a *pro*-drop language), and adjunct extraction out of islands is otherwise unacceptable. I provided a phasal account of these asymmetries based on Bošković’s (2015, 2016) Phase Collapsing and Nunes’s (2014, 2016) approach to successive cyclic movement. The account also explained an object/adjunct asymmetry Bošković noted regarding improvement under Phase Collapsing, as well as why subjects pattern with adjuncts and why there is improvement under ‘distance’.

References

- Bošković, Željko. 2005. On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59. 1–45. <https://doi.org/10.1111/j.1467-9582.2005.00118.x>
- Bošković, Željko. 2007. On the locality and motivation of move and agree: An even more minimal theory. *Linguistic Inquiry* 38. 589–644. <https://doi.org/10.1162/ling.2007.38.4.589>
- Bošković, Željko. 2013. Traces do not head islands: What can PF deletion rescue? In Yoichi Miyamoto, Daiko Takahashi, Hideki Maki, Masao Ochi, Koji Sugisaki & Asako Uchibori (eds.), *Deep insights, broad perspectives: Essays in honor of Mamoru Saito*, 56–93. Tokyo: Kaitakusha.
- Bošković, Željko. 2014. Now I’m a phase, now I’m not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45. 27–89. https://doi.org/10.1162/LING_a_00148
- Bošković, Željko. 2015. From the complex NP constraint to everything. *The Linguistic Review* 32. 603–669. <https://doi.org/10.1515/tlr-2015-0006>
- Bošković, Željko. 2016. Contextual phasehood and the ban on extraction from complements of lexical heads: When does X become a phase?. In Miyoko Yasui & Manabu Mizuguchi (eds.), *Phase theory and its consequences*, 5–39. Tokyo: Kaitakusha.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2000. Minimalist inquiries. In Roger Martin, David Michaels & Juan Uriagereka (eds.), *Step by step: Essays on minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
- Demonte, Violeta & Olga Fernández Soriano. 2005. Features in Comp and syntactic variation: The case of ‘(de)queísmo’ in Spanish. *Lingua* 115. 1063–1082. <https://doi.org/10.1016/j.lingua.2004.02.004>

- den Dikken, Marcel. 2007. Phase extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33. 1–41. <https://doi.org/10.1515/TL.2007.001>
- Gallego, Ángel. 2007. Phase theory and parametric variation. Doctoral dissertation, Universitat Autònoma de Barcelona.
- Gallego, Ángel & Juan Uriagereka. 2007. Sub-extraction from Subjects. In José Camacho, Nydia Flores-Ferrán, Liliana Sánchez, Viviane Déprez & Marí José Cabrera (eds.), *Romance Linguistics* 2006, 149–162. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.287.12gal>
- Huang, Cheng-Teh James. 1982. Logical relations in Chinese and the theory of grammar. Doctoral dissertation, MIT.
- Nunes, Jairo. 2014. Adjunct control and edge features. In Peter Kosta, Steven L. Franks, Teodora Radeva-Bork & Lilia Schürcks (eds.), *Minimalism and beyond: Radicalizing the interfaces*, 79–108. Amsterdam: John Benjamins. <https://doi.org/10.1075/lfab.11.04nun>
- Nunes, Jairo. 2016. Edge features on phase heads or moving elements? *Florida Linguistics Papers* 3. 63–80.
- Saab, Andrés. Manuscript. *Hacia una teoría de la identidad parcial en la elipsis*. Universidad de Buenos Aires. 2008.
- Uriagereka, Juan. 1988. On government. Doctoral dissertation, University of Connecticut.

On another apparent violation of the subject-island constraint in French

Guido Mensching and Franziska Werner

Seminar für Romanische Philologie, Georg-August-Universität Göttingen

This chapter addresses extractions of *wh*-marked complements of nouns out of French subject DPs into direct interrogatives – an apparent violation of the subject island constraint. We explain why some speakers of French can extract such constituents into interrogatives with complex inversion, whereas the grammaticality of other interrogative structures is clearly degraded. Our formal analysis is based on the Minimalist Program and assumes that material extracted from DPs has to pass through the DP phase-edge. In complex inversion, a structure in which the subject itself needs to move to the CP, the reordered subject DP (with the complement of N at the DP phase-edge) moves as a whole, thus giving a surface order that violates the subject island constraint only in appearance.

Keywords: French, long distance dependencies, *wh*-extraction, subject-island constraint, subject condition, generative grammar, minimalist program

1. Introduction

1.1 Aims and structure of this chapter

This contribution addresses extractions of *wh*-marked complements of nouns out of subject DPs into French direct interrogatives. In particular, we explain why some speakers of French can extract such constituents in interrogatives with so-called complex inversion, whereas the grammaticality of other interrogative structures is clearly degraded.

Extraction from DPs has been extensively discussed, by Ross (1967); Sportiche (1981), Obenauer (1984, 1994) and Chomsky (1986); Cinque (1990) and Szabolcsi (2006), among many others. Subject DPs are considered strong islands (cf. Szabolcsi & den Dikken 2003); as stated by the Subject Condition (SC), “No element may be moved out of a subject” (Müller 2011: 48). The SC is usually considered to be

universal and has consequently also been observed for French:¹ whereas extraction out of direct object DPs is grammatical with ‘genitive *de*’ (cf. Sportiche 1981; Obenauer 1984; Pollock 1989; Tellier 1990; Godard 1992; Mensching 2019), as shown in (1a), extraction out of a subject DP is ungrammatical, as shown in (1b).

- (1) a. *De quel livre connais-tu la fin* __ ?
of which book know=you the end
“Of which book do you know the end __?”² (Sportiche 1981: 224)
- b. [?]*De quel linguiste est-ce que les parents* __ *ont déménagé*
of which linguist is-it that the parents have moved
à *Chartres* ?
to Chartres
“Of which linguist did the parents __ move to Chartres?”
(Tellier 1991: 89)

However, some exceptions have been noted for French. Tellier (1990: 90–91) observes that extraction is possible from subjects of copular and passive structures, and similar SC violations are found with unaccusative verbs.³ According to Tellier, this is because these subjects are ‘derived subjects’: if the SC refers to Spec,VP, the structures at issue are not SC violations. Another case observed by Tellier (1990: 92–97; 1991) concerns the relativizer *dont* ‘of which’, which is extractable even from non-derived subjects.⁴ For Tellier, this ultimately follows from *dont*’s status as a complementizer (C°).⁵ More recently, Heck (2008, 2009) has presented another solution (see § 3.1 for details). Importantly, in Heck’s account, *dont* moves together with the subject, so here the SC violation is only apparent.

In this chapter, we report on yet another exception to the SC that is found in at least some speakers of French. These speakers indicate a surprising increase in acceptability for extractions of *wh*-marked nominal complements out of subject DPs into direct interrogatives with complex inversion (CI) as compared to the same sentences using the interrogative *est-ce que* structure (ESQ):

-
1. For another view, see Stepanov (2007).
 2. For illustrative purposes, we prefer rather literal translations that do not pretend to be grammatical in English. The extraction site is marked by __ both in the original sentence and in the translation.
 3. For similar observations concerning other languages, cf., e.g., Uriagereka (1988: 118) and Chomsky (2008: 153–154), quoted in Spyropoulos & Stamatogiannis (2011).
 4. See Example (12) in § 3.1.
 5. Tellier’s explanation is based on subjacency and Chomsky’s (1986) *Barriers* framework.

- (2) a. ^{??6}*De quel linguiste est-ce que les parents __ ont déménagé*
 of which linguist is-it that the parents have moved
à Chartres ?
 to Chartres
 “Of which linguist did the parents __ move to Chartres?”
 complex inversion (CI)
- b. [?]*De quel linguiste les parents __ ont-ils déménagé à Chartres ?*
 of which linguist the parents have=they moved to Chartres
 “Of which linguist did the parents __ move to Chartres?”
est-ce que (ESQ)

In § 4, we argue that the high degree of grammaticality in examples such as (2b) is another case of an SC violation that is only apparent: similarly to Heck’s (2008, 2009) interpretation of the *dont* cases, we show that cases such as (2b) are best to be analyzed as the result of the whole subject DP being extracted.

This chapter is organized as follows: § 1.2 briefly describes French *wh*-question types. In § 2, we present more data on the (apparent) SC violation at issue. Section 3 summarizes some assumptions that are necessary to explain our data, in particular concerning the extraction of N-complements from direct objects in French and concerning CI. In § 4, we argue that examples like (2b) can be explained roughly as follows: the *wh*-marked complement moves to the left edge of the DP, followed by movement of the whole subject DP to the CP. Section 5 briefly provides a summary and an outlook.

Our explanation is based on the Minimalist Program, following Chomsky (2001 et seqq.), so we assume phase theory as well as movement licensed by Agree and triggered by movement-inducing features. More details are found in § 3 and § 4.

Finally, some additional words on the scope and limitations of this article are in order: (i) Since the SC has been observed not to hold for “subjects” of copular and unaccusative structures, we concentrate on transitive verbs,⁷ (ii) French speakers differ from each other in their judgments of sentence pairs like those in (2a),(b). We interpret this as being related to speakers’ having different grammars. We particularly focus on one of these types of grammars, represented by one group of speakers. A further investigation of other types of French grammars as well as the nature of this phenomenon is beyond the scope of this chapter, but we reflect briefly on this issue in § 5.

6. Note the slightly different judgment (as ?*) by Tellier in Example (1b).

7. Unergative verbs are less appropriate for testing SC-violations because some unergative verbs show ‘unaccusativization’. See Mensching & Weingart (2016, in particular note 43 on p. 319) for a brief overview.

1.2 Types of direct *wh*-interrogatives in French

Modern Standard French has the five *wh*-question types illustrated below. All types convey the same meaning (“When did Marie/she arrive?” in (3) to (7)).⁸

- | | | |
|-----|---|-------------------------|
| (3) | <i>Marie est arrivée quand ?</i>
Marie is arrived when | <i>wh</i> -in-situ |
| (4) | <i>Quand est arrivée Marie ?</i>
when is arrived Marie | stylistic inversion |
| (5) | <i>Quand est-elle arrivée ?</i>
when is=she arrived | clitic inversion |
| (6) | <i>Quand Marie est-elle arrivée ?</i>
when Marie is=she arrived | complex inversion (CI) |
| (7) | <i>Quand est-ce que Marie est arrivée ?</i>
when is-it that Marie is arrived | <i>est-ce que</i> (ESQ) |

Wh-in-situ questions like those in (3) do not involve overt *wh*-movement and thus do not show SC effects. In questions with ‘stylistic inversion’ (illustrated in (4)), the subject appears after the whole verbal complex (Aux+V). But they are not grammatical with unergative and transitive verbs⁹ and are thus also of no interest here. Equally irrelevant is the ‘clitic inversion’ shown in (5), because the subject is necessarily a clitic.

Thus, we can concentrate on the remaining two types. In CI-questions like in (6), the subject is placed in the preverbal position and is doubled by a subject clitic following the finite verb, whereas the *wh*-constituent precedes the subject (see § 3 for a formal analysis). ESQ-questions like in (7) contain *est-ce que* [*lit.* “is it that”] appearing between the moved *wh*-item and the subject. There are basically two lines of analysis for this type of question in generative grammar: *est-ce que* is either a question particle in C° (see, e.g., Rooryck 1994; Kellert 2017) or it can be analyzed compositionally, as in Munaro & Pollock (2005).¹⁰

8. For details see, e.g., Jones (1996: 464–467).

9. For these and other restrictions, see Kayne & Pollock (1978); Jones (1996); Bonami & Godard (2001).

10. Their analysis involves a ‘copular phrase’ ([_{CopP} [_{Cop°} est] [_{SmallCl} ce, ___]]) in the CP with one operator phrase (Op₂P) in a higher position and another one (Op₁P) in a lower position, also belonging to the left periphery. Op₁° contains *que*, whereas Spec,Op₂P is the final landing site of the *wh*-element, which passes through the empty position in CopP.

2. The data

The observation that SC violations can be tolerated with CI goes back to a small survey of speakers conducted in Toulouse in 2014.¹¹ Apart from Example (2b),¹² we note the following sentences, which yielded even higher grammaticality scores:¹³

- (8) a. *De quel linguiste les parents __ ont-ils acheté une maison*
 of which linguist the parents have=they bought a house
à Chartres ?
 at Chartres
 “Of which linguist did the parents __ buy a house in Chartres?”
- b. *De quel auteur plusieurs traductions __ ont-elles gagné des*
 of which author several translations have=they won INDEF.ART
prix internationaux ?
 prizes international
 “Of which author did several translations __ win international prizes?”

A recent informal replication of our findings with three speakers of different dialectal backgrounds confirmed that SC violations tend to be tolerated more readily with CI than with ESQ. We later conducted an online inquiry on a larger scale.¹⁴ The sentences yielded rather low average grammaticality scores (CI: 4.06, ESQ: 3.61, with 10 = grammatical),¹⁵ but the overall result shows that extraction from subjects is marginally significantly better ($p < 0.1$) with CI than with ESQ. For our purposes, we can distinguish three types of speakers (see Appendix):¹⁶

11. 18 native speakers of French, 11 relative and interrogative clauses with extractions from subjects. The task was to rate the grammaticality on a scale from 0 (ungrammatical) to 3 (fully grammatical). We thank Injoo Choi-Jonin for permitting us to distribute the questionnaire among her students.

12. Average grammaticality on the scale mentioned in note 11: 1.1 for (2a), 2.0 for (2b).

13. 2.5 for (8a) and 2.8 for (8b).

14. The 22 participants, mostly from different parts of France, were presented with 14 randomized interrogatives with transitive verbs (7 with ESQ and 7 with CI, see (9) and (10)), and distractors. Participants were asked to rate the sentences on a Likert scale: 1 to 10 (with 1 being ‘not acceptable’ and 10 ‘completely acceptable’).

15. The sentences containing complex DPs with *quel* ‘which’ were often judged better than those containing *qui* ‘who’, a phenomenon (possibly related to D-linking) that we cannot discuss in the present article.

16. The results of the small 2014 Toulouse survey (see note 11) only reflected ratings by groups 1 (6 speakers, i.e., 35.3%) and 2 (11 speakers, i.e., 64.7%). Note that one of the 18 participants only rated the relative clauses and not the interrogative clauses.

1. speakers who judge ESQ and CI equally (un)acceptable: (10 speakers out of 22, i.e., 45.45%),
2. speakers who prefer CI over ESQ (8 speakers out of 22, i.e., 36.36%, and
3. speakers who prefer ESQ over CI (4 speakers out of 22, i.e., 18.18%).

Going beyond statistical significance, it is remarkable that, with all speakers taken together, of the overall 308 items, 62 items (i.e., approx. 20%) are rated higher than 7, which would be unexpected if subject extractions were really ungrammatical for all speakers of French. Of these 62 cases, 28 are ESQ-questions, whereas 34 are CI-questions, which again confirms the preference for CI in *wh*-extractions from subject DPs. Even more strikingly, if we consider only those sentences that were given the highest score (i.e., 10 on the Likert scale), only 4 correspond to ESQ¹⁷ and 16 to CI.¹⁸

While we briefly return to these results (and, in particular to the speakers of type 3) in § 5, the rest of the chapter focuses on speakers of type 2. In (9) and (10) we reproduce the results of one speaker (speaker 16). This speaker was chosen mainly for illustrative purposes, as, within group 2, he shows the strongest deviance between the grammaticality values for CI and ESQ.¹⁹

- (9) a. [7] *De qui les idées __ ont-elles inspiré tes réflexions ?*
of whom the ideas have=they inspired your reflections
“Of whom have the ideas __ inspired your reflections?”
- b. [9] *De quel collaborateur l' absence __ a-t-elle perturbé l' organisation de la dernière réunion ?*
of which coworker the absence has=it disturbed the organization of the last meeting
“Of which coworker did the absence __ disturb the organization of the last meeting?”
- c. [10] *De qui le fils __ a-t-il vendu sa nouvelle voiture ?*
of whom the son has=he sold his new car
“Of whom did the son __ sell his new car?”
- d. [10] *De quelle patiente le mari __ cherche-t-il du travail ?*
of which patient the husband looks.for=he of.the work
“Of which patient does the husband __ look for work?”

17. One judgment by speaker 7 and 3 by speaker 12.

18. Speakers 7, 9, 20, and 21: 1 item each; speaker 12: 2 items, speakers 3 and 22: 3 items each; speaker 16: 4 items. The latter speaker is the one whose data we will examine more closely below (see the examples in (9) and (10)).

19. The numbers between square brackets refer to the rating, see fn. 14.

- e. [10] *De qui la chef __ a-t-elle créé une nouvelle filiale*
 of whom the boss has=she established a new branch
à Marseille ?
 at Marseille
 “Of whom did the boss __ establish a new branch in Marseille?”
- f. [5] *De quel ami les parents __ ont-ils ouvert une*
 of which friend the parents have=they opened a
boulangerie à Paris ?
 bakery at Paris
 “Of which friend did the parents __ open a bakery in Paris?”
- g. [10] *De quelle créatrice de mode les vêtements __ émerveillent-ils*
 of which fashion.designer the clothes captivate=they
ta sœur ?
 your sister
 “Of which fashion designer do the clothes __ captivate your sister?”
- (10) a. [1] *De qui est-ce que les parents __ ont rénové la maison*
 of whom is-it that the parents have renovated the house
de tes grands-parents ?
 of your grandparents
 “Of whom did the parents renovate the house of your grandparents?”
- b. [5] *De quelle artiste est-ce que la vidéo __ a eu un*
 of which artist is-it that the video has had a
joli succès ?
 great success
 “Of which artist has the video __ been a great success?”
- c. [1] *De qui est-ce que les élèves __ ont rédigé une dissertation*
 of whom is-it that the pupils have edited a dissertation
sur Victor Hugo ?
 on Victor Hugo
 “Of whom did the pupils __ write a dissertation on Victor Hugo?”
- d. [3] *De quels auteurs est-ce que les textes __ ont un dénominateur*
 of which authors is-it that the texts have a denominator
commun ?
 common
 “Of which authors did the texts __ have a common denominator?”
- e. [4] *De qui est-ce que les enfants __ vont bâtir leur*
 of whom is-it that the children go build their
propre maison ?
 own house
 “Of whom will the children __ build their own house?”

- f. [4] *De quel collègue est-ce que l' épouse__ a préparé*
 of which colleague is-it that the wife has prepared
un bon repas ?
 a delicious meal
 "Of which colleague did the wife __ prepare a delicious meal?"
- g. [1] *De quel footballeur est-ce que les fans__ ont chanté l'*
 of which football.player is-it that the fans have chanted the
Hymne national ?
 anthem national
 "Of which football player did the fans __ chant the national anthem?"

The speaker rated the *wh*-questions with CI in (9) between [5] and [10]: four out of seven sentences, (9c),(d),(e),(g), were perfectly grammatical; (9b) is nearly perfect with [9], and only (9a) and (9f) got lower scores, with [7] and [5], respectively. By contrast, the speaker rated the ESQ-questions in (10) between [1] and [5]. Three out of seven, (10a),(c),(g), seem to be totally ungrammatical for this speaker, and only (10b) was rated with [5].

3. Theoretical background on extraction from DP and complex inversion in French

3.1 Extraction from DPs in French

The extraction of both relative pronouns and *wh*-items out of direct objects is grammatical in French but subject to several constraints (cf., among others, Grosu 1974; Tellier 1990; Sportiche 1981; Obenauer 1984; Pollock 1989; Godard 1992; Kolliakou 1999). First, such extractions are only licensed for constituents introduced by *de*, see (11a), (b) vs (11c); second, only complements of N are extractable, but not adjuncts, see (11a), (b) vs (11c),(d).

- (11) a. [_{PP} *De quel livre*]_i *connais-tu* [_{DP} *la fin* t_i]_?
 of which book know=you the end
 "Of which book do you know the end __?" (cf. Sportiche 1981: 224)
- b. [_{PP} *De qui*]_i *avez-vous vu* [_{DP} *une amie* t_i]
 of whom have=you seen a friend
 "Of whom did you see a friend __?" (cf. Grosu 1974: 312, fn. 3)
- c. * [_{PP} *Sur qui*]_i *as-tu lu* [_{DP} *le livre* t_i]_?
 on whom have=you read the book
 "On whom have you read the book __?" (cf. Mensching 2019)
- d. * [_{PP} *De quel pays*]_i *as-tu mangé* [_{DP} *des bananes* t_i]_?
 of which country have=you eaten INDEF.ART bananas
 "From which country did you eat bananas?"

According to the literature, extractable N-complements such as those in (11a), (b), which can either express the agent, theme, or possessor of (the action expressed by) the noun, are what have been dubbed ‘genitive PPs’. More recent generative literature argues that genitive case assignment actually plays a role in the extraction mechanism; see, for example, Gutiérrez-Bravo (2001) for similar extractions in Spanish. Gutiérrez-Bravo considers DPs to be phases²⁰ and assumes that only constituents that have both a genitive feature and a [*wh*]-feature can be attracted to the phase-edge of the direct object DP as a necessary intermediate step (see Cinque 2014 for a similar view).

Mensching (2019) adapts Gutiérrez-Bravo’s approach (which, apart from phase theory, is still based on Chomsky 1995)²¹ to more recent Minimalism: extracted constituents such as those in (11a), (b) are taken to be DPs²² (and not PPs) that contain both valued phi-features ([*vφ*]) and valued operator features²³ ([*vOp*]), which can take the values [*wh*] and [*rel(ative)*]. In addition, they have an unvalued case feature ([*uCase*]), which identifies them as arguments and make the DPs visible to the probe (so adjuncts are not found). Agree takes place between the D-head of the higher DP, in which the extractee is embedded, and the extractee itself (i.e., the complement of the noun). Mensching (2019) assumes that D° with [*uφ*]²⁴ assigns genitive case under Agree (cf. Radford 2004: 368–369; Rappaport 2006, among others). In addition, a French D-head with [*uφ*] can optionally have unvalued operator features ([*uOp*]) together with an [*EPP*]-feature; this option becomes relevant in the case of extraction. Figure 1 illustrates the probing mechanism and how the extractee is moved to the DP phase-edge.²⁵

20. See also Svenonius (2004); Chomsky (2008); Heck (2008, 2009); Cinque (2014), among many others.

21. He assumes an AgrGen[itive] projection beneath D. In order to attract the *wh*-constituent to the phase border, D° has a strong *wh*-feature. The covert AgrGen° head has a [GEN]-feature, which is adjoined to D and then attracts the PP with [GEN] to Spec,DP. The attracted PP must also bear a [*wh*]-feature, which checks D°’s [*wh*]-feature.

22. Or case phrases (KPs). See Mensching (2019) for discussion.

23. Cf. Radford (2004: 419ff.).

24. A second phi-probe, which regulates agreement in gender and number between D° and N°, is ignored here; see Mensching (2019) for a more complete derivation.

25. An anonymous reviewer remarks that this assumption raises the issue of pied-piping constructions of the type *Susan, the picture of whom John likes, ...*, which also exist in French, and where there is no indication that the *wh*-element moves overtly (‘massive pied-piping’ according to Safir 1986). For an explanation, see Heck (2008: 300), who assumes “that appositive relative clauses and matrix interrogatives can employ a mechanism that prevents phrase and phase boundaries that are pied-piped from causing violations of LA [i.e., Local Agree, GM and FW]

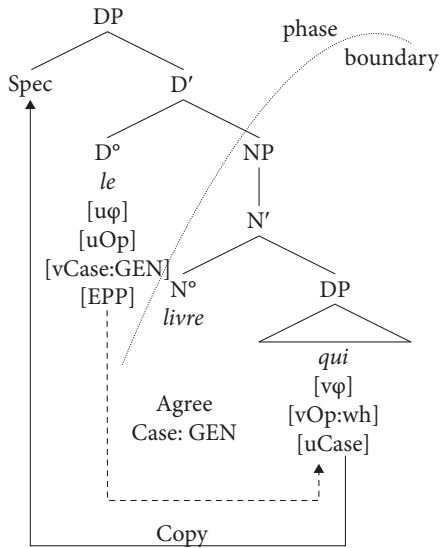


Figure 1. Movement of the N complement to the DP phase-edge:
le livre (de) qui ==> *(de) qui le livre*

From the phase-edge, the *wh*-constituent can be extracted to C° by an appropriate probe. Mensching (2019) assumes that *de* is inserted post-syntactically as the spell-out of [GEN].

The idea that extractions from a DP must pass through its phase-edge also underlies Heck's (2008, 2009) analysis of the relative item *dont* "of which". The grammatical Example (12) should be barred because of the SC:

- (12) *un homme dont le comportement devient drôle*
 a man of.which the behavior becomes weird
 "a man of which the behavior becomes weird"
 (meaning "a man whose behavior becomes weird")

or the PIC although, apparently, Agree between [$*WH^*$] and [WH] is established across them. The idea is that probe and goal are in fact sufficiently close to each other in order to avoid fatal violations of these constraints". For details, see Heck (2008: Chapter 5). Massive pied-piping also exists as an option in French *wh*-questions, see (i):

- (i) *Les parents de qui ont déménagé à Chartres ?*
 the parents of whom have moved to Chartres
 "The parents of whom (i.e., whose parents) have moved to Chartres?"

As no extraction is involved here, this construction is not considered in this chapter.

Heck (2008: 96) analyzes this construction as *wh*-movement of *dont* to the DP phase-edge (“secondary *wh*-movement within pied-piped constituents”) followed by movement of the whole DP:

- (13) *Un homme* [_{CP} [_{DP} *dont*₂ *le comportement* *t*₂]₃ [_{TP} *t*₃ *devient* [_{VP} *t*₃ *drôle*]]].
(cf. Heck 2008: 96)

As the whole subject (including *dont*) moves to Spec,CP, there is no SC violation.

3.2 Complex inversion in French

Here, we first briefly summarize Rizzi & Roberts’ (1989) analysis of French CI. In a second step, we sketch what a minimalist version could look like. CI is illustrated again in (14):

- (14) *Quel livre Jean a-t-il lu ___?*
which book Jean has=he read
“Which book did Jean read?”
(Rizzi & Roberts 1989: 1)

Rizzi & Roberts assume that the subject (*Jean*) is the argumental subject, base-generated in Spec,VP whereas the clitic *il* is a double, which is generated in Spec,IP and incorporates into C° together with I°. The *wh*-expression (*quel livre*) moves to Spec,CP, whereas the subject is adjoined to C’ (see Figure 2):

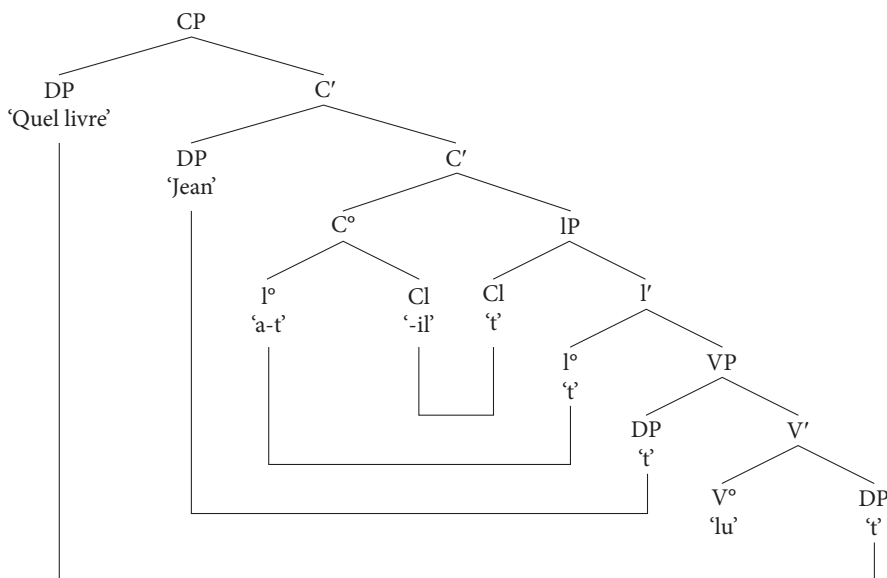


Figure 2. Rizzi & Roberts’ (1989) analysis, tree structure by Lalande (1997: 101)

In order to adapt this approach to the Minimalist Program, we have to resolve at least two problems: first, the clitic needs an explanation, and, second, subject movement to the CP has to be motivated. The two issues can be resolved together if we assume unvalued phi-features in C° that probe the subject: the clitic can then be considered as the spell-out of this phi-probe after its valuation.²⁶ In addition, since we have a *wh*-construction, C° also needs an operator probe ([uOP], see § 3.1). Both probes need an [EPP]-feature that ultimately triggers the movements. For illustration, see Figure 3:

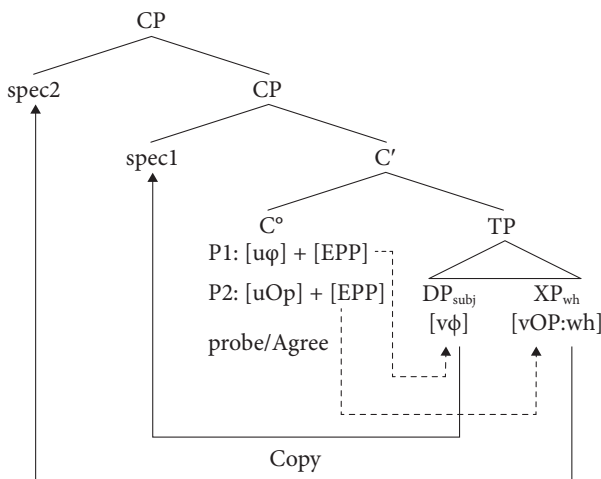


Figure 3. Minimalist analysis of French complex inversion (P= probe)

The two probes have to probe in the order indicated in Figure 3. Probe 1 (P1) finds the subject, and the [EPP]-feature will attract it to Spec,CP (spec1). Probe 2 (P2) finds the *wh*-item, which is attracted to an outer specifier of CP (spec2).

26. As an anonymous reviewer remarks, this amounts to treating complex inversion as a special case of agreement, which is a traditional approach to subject clitic inversion generally known as *conjugaison interrogative* (see Pollock 2006 and Roberts 2010 for motivation). Also note that, under standard minimalist assumptions, to be visible, the subject needs [uCase]. However, this feature has already been valued as nominative by T° . The issue of why P1 can still see the subject and no freezing occurs needs some further research. We think that a possible solution might be related to the fact that, according to Chomsky (2008), the features of T are inherited from C. One track that might be examined is whether the clitic and the verbal inflection are both spell-outs of the same probe, which is also responsible for nominative assignment, so the subject remains visible also for the C head.

4. Analysis

We now analyze the data presented in § 2 and answer the question as to why *wh*-extractions from subjects are (more) grammatical with CI. Based on the assumptions in § 3.2, a sentence such as (15a) (repeated from (9c)) should have the structure in (15b) (post-syntactically inserted material is boldfaced):

- (15) a. *De qui le fils a-t-il vendu sa nouvelle voiture ?*
 of whom the son has=he sold his new car
 “Of whom did the son ___ sell his new car?”
- b. [_{CP} [_{DP} ***De qui***]_i [_{CP} [_{DP} *t_i le fils* *t_i*]_j [_{C°} *a_k-t-il*] [_{TP} *t_k* [_{VP} *t_j vendu*_l [_{VP} *t_l sa nouvelle voiture*]]]]]?

In (15b), the outer specifier of CP is occupied by *de qui*, the complement of the noun *fils*. The remnant of the subject DP has moved to the inner specifier of CP, following the derivation of CI-structures in Figure 3 of § 3.2. However, the derivation in (15b) should be barred due to the SC. Somehow inverting Rizzi & Roberts’ (1989) approach (see § 3.2), in order to save this derivation, we could assume that the clitic is the argumental subject in Spec,vP and that the full DP is in Spec,TP. We could then stipulate that the SC is not valid for the DP, not being a ‘true’ (argumental) subject. However, we would probably run into theta-theoretic problems, and this solution rather seems ad hoc and would force us to abandon the minimalist interpretation presented in § 3.2, which has the advantage of motivating the presence of the clitic.

We think that the solution is much easier and follows from the theories presented in § 3. The characteristic property of CI, which distinguishes this construction from all other French interrogatives, is that the subject itself moves to the CP. In § 3.2, we formalized this with [φ]+[EPP] in C°. The same head also contains [uOp], which has to agree with [vOp] of *qui* in (15a). But the *wh*-item must have previously moved to the DP phase-edge in order to be visible to a probe (see § 3.1). Therefore, similarly to Heck’s (2008, 2009) approach for relative *dont*, we adopt the following solution: the whole subject DP moves (after ‘secondary *wh*-movement’ within the DP, see (16)), as shown in (17):²⁷

27. An anonymous reviewer objects to this analysis by saying that CI-questions such as (i) with the subject itself situated in the CP are ungrammatical, but are structurally similar to what we assume here:

- (i) **Qui a-t-il lu le livre ?*
 who has-he read the book
 “Who has read the book?”

However, according to Rizzi & Roberts (1989: 15, following Kayne 1983), the correct description of the case in (i) is that “the construction [i.e., complex inversion] does not allow questioning of the subject itself”. Instead, in our construction, only a part of the subject is questioned.

(16) $[_{DP} le [_{NP} fils [_{DP} de qui]]] \Rightarrow [_{DP} [_{DP} de qui]_i [_{DP} le [_{NP} fils t_i]]]$

In the end, both the subject and the *wh*-constituent are in the CP (but in only one specifier, differently from Figure 3 above). This yields the right surface order:²⁸

(17) $[_{CP} [_{DP} [_{DP} de qui]_k [_{DP} le fils t_k]_i [_{C^\circ} a_k-t-il] [_{TP} t_j t_k [_{VP} t_j vendue] [_{VP} t_i sa nouvelle voiture]]]]]$

Hence, the data presented in § 2 are no violations of the SC, because nothing is actually extracted. Our analysis is shown in more detail in Figure 4:

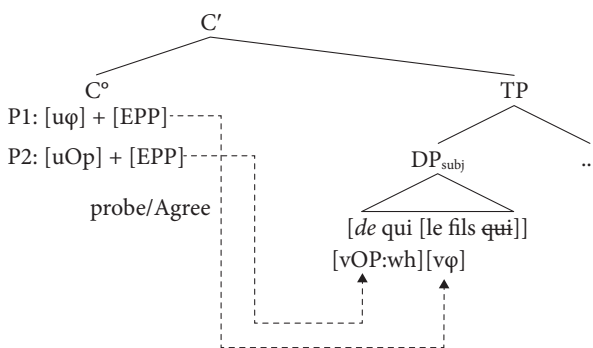


Figure 4. Pseudo-extraction out of a subject with complex inversion

A potential problem is that probe 1 (P1) attracts the whole DP to Spec,CP, thus preventing probe 2 (P2) from probing. Possible solutions might be: (a) the feature of P2 is checked as a kind of ‘free rider’, (b) P1 and P2 probe simultaneously, or (c) the probing order is inverse (i.e., P2 probes before P1).²⁹ For the purposes of the present contribution, we prefer to adopt solution (b).³⁰ In any case, the final structure should be as shown in Figure 5, with no extraction from a subject position.

Another problem may be that one [EPP]-feature remains unchecked. We assume here that both [EPP]-features can actually be checked because both items move, albeit in one step instead of two. If this is undesirable, one may think of

28. Recall from § 3.2 that the clitic is the spell-out of the phi-probe in C° after Agree with the subject.

29. If this is correct, then ‘regular’ complex inversion (like in (14)) would be a case of ‘tucking in’ (Richards 2001) of the subject, hence the probes in Figure 3 (and thus also in Figure 4) must work in inverse order. In this case, P2 would actually find the goal and perform the operation Agree; however, movement would not occur because of the SC.

30. As an anonymous reviewer points out, a Determinacy approach (Chomsky et al. 2019) might be interesting for resolving these problems in future work.

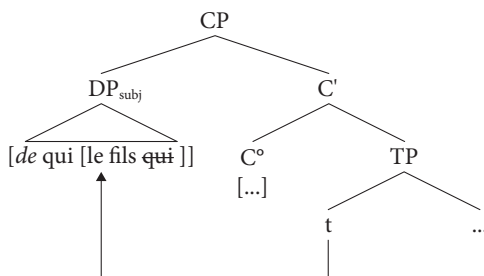


Figure 5. Final structure of pseudo-extraction out of a subject DP

further movement of the *wh*-phrase to an outer CP-specifier. Note that this would probably not count as an SC violation because Spec,CP is not a subject position.³¹

Finally, note that there might be independent evidence for this analysis, something that needs to be examined more closely in future research. A small additional informal inquiry has shown that, for speakers who accept extraction from subjects in interrogatives with CI, the insertion of an adverb between *de qui* and the DP heavily degrades the sentence:³²

- (18) ?**De qui malheureusement les parents __ ont-ils moved*
of whom unfortunately the parents have=they déménagé
to Chartres
à Chartres ?
‘Of whom have the parents __ unfortunately moved to Chartres?’

Examples such as (18) seem to be more acceptable with items such as *donc* ‘so’ or *alors* ‘then’ instead of *malheureusement*, but this fact does not contradict our analysis, as these items, being discourse markers, can be seen as parentheticals, that is, as constituents that “seem to be linearly integrated in, but structurally independent from, their host” (Dehé & Kavalova 2007: 4, cf., among many others, Haegeman’s 1988 interpretation of parentheticals as ‘orphans’).³³

31. However, this would be countercyclic (if it happens at the CP-level) and would violate freezing. Thanks to an anonymous reviewer for pointing this out.

32. The adverb seems to be more acceptable, however, when it appears after *les parents* or after the verb+clitic sequence.

33. An anonymous reviewer finds (18) much better than ?*, also observing that “[t]he reason ... [(18)] may be a little worse for some speakers is the type of adverb that is used, and the type of intonation that licenses the interpolation of the adverb here”. We think that this might indicate that, for this reviewer, *malheureusement* is a parenthetical (note that the English *unfortunately* is mentioned as an example of an adverb used as a parenthetical in Dehé & Kavalova 2007: 2). We think that this is confirmed by examples of the type in (i), which the reviewer

5. Conclusions and outlook

We have detected a French interrogative structure that seems to be a violation of the SC: some speakers of French can extract DPs from subjects in interrogatives with the configuration known as ‘complex inversion’ (CI), which appears (on first sight) to be an extraction out of a subject DP. We have explained this by assuming ‘secondary *wh*-movement in the DP’ (Heck 2008, 2009) followed by movement of the whole subject DP to Spec,CP, as required in CI. If our analysis is correct, there is actually no SC violation in the sentences at issue. Similar (but not identical) cases are those described by Heck’s “Edge Generalization” (2008: 88).

We have been concerned with a particular grammar of some speakers (let us call it grammar A). These speakers reject similar structures with ESQ-questions. Other speakers, who represent another grammar of French (grammar B), find such cases more or less acceptable with the ESQ-construction.³⁴ Grammar B has not been part of this chapter and is currently being examined by our research group. We can nevertheless provide some initial ideas on why grammar B-speakers may accept SC violations in ESQ-questions. A first idea is that this is a ‘grammatical illusion’ in the sense of Haider (2011). In this case, the ‘illusion’ is caused by the presence of the copula (*est*) within the *est-ce que* string, bearing in mind that French copula constructions do not fall under the SC (see § 1). This might also explain why even the grammar A-speaker whose data we examined more closely in § 2, assigns some degree of grammaticality to some of the ESQ-examples in (10). A second idea is that the two lines of analysis existing for the French ESQ-construction (see, again, § 1) are both correct in a speaker-dependent way: in particular, it might be the case that grammar B-speakers generate ESQ-questions compositionally along the lines of Munaro & Pollock’s (2005) analysis and can extract from higher (intermediate)

regards as counterexamples to (18) and, thus, to our theory. Instead, we think that they are referring to parentheticals, as their use of the commas shows:

- (i) *De qui, hier soir, le fils __ a-t-il vendu sa nouvelle voiture ?*
 of whom yesterday evening the son has=he sold his new car
 “Of whom, yesterday evening, has the son __ sold his new car,” meaning “Whose son did sell his car yesterday?”

The reviewer also says that our theory cannot explain why the intercalation of an adverbial between the *wh*-item and its complement in massive pied-pied constructions like those mentioned in note 25 “is strongly ungrammatical (**)”. We think the reason is that parentheticals cannot intervene between a head noun and its ‘genitive’ complement.

34. The nature of this variation is unclear for now. The results shown in the Appendix suggest that the relevant variables may be age and/or location, but much more data and a different method would be necessary to further elucidate this point.

landing sites of the subject DP (outside the TP-vP domain and thus possibly not counting for the SC). Conversely, for grammar A-speakers, *est-ce que* is possibly a question particle and there are no such higher landing sites of the subject. However, as we said, we must leave this for future research.

Acknowledgements

We thank Elodie Winckel for her help with the planning of the survey described in § 2 and the statistical evaluation of the results.

Funding

Our research was funded by the Deutsche Forschungsgemeinschaft (“Long-Distance-Dependencies in French: Comparative Analyses [HPSG and the Minimalist Program]”); principle investigators: Stefan Müller and Guido Mensching, researchers: Elodie Winckel and Franziska Werner.

References

- Bonami, Olivier & Danièle Godard. 2001. Inversion du sujet: Constituance et ordre des mots. In Jean-Marie Marandin (ed.), *Cahier Jean-Claude Milner*, 117–174. Lagrasse: Verdier.
- Chomsky, Noam. 1986. *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1995. *The minimalist program*. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. Derivation by phase. In Michael J. Kenstowicz (ed.), *Ken Hale: A life in language*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2008. On phases. In Robert Freidin, Carlos Peregrín Otero & Maria 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>
- Chomsky, Noam, Ángel J. Gallego & Dennis Ott. 2019. Generative grammar and the faculty of language: Insights, questions, and challenges. In Ángel J. Gallego & Dennis Ott (eds.), *Generative syntax: Questions, crossroads, and challenges* (special issue of *Catalan Journal of Linguistics*), 229–261. Barcelona: Universitat de Barcelona, Servei de Publicacions.
<https://doi.org/10.5565/rev/catjl.288> (8 July, 2020).
- Cinque, Guglielmo. 1990. *Types of \bar{A} -dependencies*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 2014. Extraction from DP in Italian revisited. In Enoch Oladé Aboh, Maria Teresa Guasti & Ian Roberts (eds.), *Locality*, 86–103. New York: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199945269.003.0004>
- Dehé, Nicole & Yordanka Kavalova. 2007. Parentheticals: An introduction. In Nicole Dehé & Yordanka Kavalova (eds.), *Parentheticals*, 1–22. Amsterdam: John Benjamins.
<https://doi.org/10.1075/la.106.03deh>

- Godard, Danièle. 1992. Extraction out of NP in French. *Natural Language & Linguistic Theory* 10. 233–277. <https://doi.org/10.1007/BF00133813>
- Grosu, Alexander. 1974. On the nature of the left branch condition. *Linguistic Inquiry* 5(2). 308–319.
- Gutiérrez-Bravo, Rodrigo. 2001. Phases, case and accessibility: The case of extraction from DP in Spanish. In James McCloskey (ed.), *Syntax and semantics at Santa Cruz* 3, 104–118. Santa Cruz, CA: Linguistics Research Center, University of California.
- Haegeman, Liliane. 1988. Parenthetical adverbials: The Radical orphanage approach. In Shuki Chiba, Akira Ogawa, Yasuaki Fuiwara, Norio Yamada, Osama Koma & Takao Yagi (eds.), *Aspects of modern linguistics: Papers presented to Masatomo Ukaji on his 60th birthday*, 232–254. Tokyo: Kaitakushi.
- Haider, Hubert. 2011. Grammatische Illusionen. *Zeitschrift für Sprachwissenschaft* 30. 223–257.
- Heck, Fabian. 2008. *On pied-piping: Wh-movement and beyond*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110211467>
- Heck, Fabian. 2009. On certain properties of pied-piping. *Linguistic Inquiry* 40(1). 75–111. <https://doi.org/10.1162/ling.2009.40.1.75>
- Jones, Michael Allan. 1996. *Foundations of French syntax*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511620591>
- Kayne, Richard S. 1983. Chains, categories external to S, and French complex inversion. *Natural Language & Linguistic Theory* 1. 107–139. <https://doi.org/10.1007/BF00210378>
- Kayne, Richard S. & Jean-Yves Pollock. 1978. Stylistic inversion, successive cyclicity, and move NP in French. *Linguistic Inquiry* 9(4). 595–621.
- Kellert, Olga. 2017. Interrogatives. In Andreas Dufter & Elisabeth Stark (eds.), *Manual of Romance morphosyntax and syntax*, 569–603. Berlin: de Gruyter. <https://doi.org/10.1515/9783110377088-016>
- Kolliakou, Dimitra. 1999. De-phrase extractability and individual/property denotation. *Natural Language & Linguistic Theory* 17. 713–781. <https://doi.org/10.1023/A:1006256028441>
- Lalonde, Jean Yves. 1997. *Verbstellung im Deutschen und Französischen: Unter Anwendung eines CAD-basierten Expertensystems*. Tübingen: Niemeyer. <https://doi.org/10.1515/9783110959987>
- Mensching, Guido. 2019. Extraction from DP in French: A minimalist approach. In Ludovico Franco, Mihaela Moreno & Matthew Reeve (eds.), *Agreement, case and locality in the nominal and verbal domains*, 249–277. Berlin: Language Science Press.
- Mensching, Guido & Anja Weingart. 2016. The null subject parameter and the lexicon in minimalist syntax. In Luis Javier Eguren Gutiérrez, Olga Fernández-Soriano & Amaya Mendikoetxea (eds.), *Rethinking parameters*, 291–328. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780190461737.003.0010>
- Müller, Gereon. 2011. *Constraints on displacement: A phase-based approach*. Amsterdam: John Benjamins. <https://doi.org/10.1075/lfab.7>
- 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.
- Obenauer, Hans-Georg. 1984. On the identification of empty categories. *The Linguistic Review* 4. 153–202.
- Obenauer, Hans-Georg. 1994. *Aspects de la syntaxe A-barre*. Doctoral dissertation, Université de Paris VIII.
- Pollock, Jean-Yves. 1989. Verb movement, universal grammar, and the structure of IP. *Linguistic Inquiry* 20(3). 365–424.

- Pollock, Jean-Yves. 2006. Subject clitics and complex inversion. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, Volume 4, 601–659. Malden, MA: Blackwell. <https://doi.org/10.1002/9780470996591.ch67>
- Radford, Andrew. 2004. *Minimalist syntax: Exploring the structure of English*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511811319>
- Rappaport, Gilbert C. 2006. The Syntax of possessors in the nominal phrase: Drawing the lines and deriving the forms. In Ji-yung Kim, Yuri A. Lander & Barbara H. Partee (eds.), *Possessives and beyond: Semantics and syntax*, 243–261. Amherst, MA: GLSA Publications.
- Richards, Norvin. 2001. *Movement in language: Interactions and architectures*. Oxford: Oxford University Press.
- Rizzi, Luigi & Ian Roberts. 1989. Complex inversion in French. *Probus* 1(1). 1–30. <https://doi.org/10.1515/prbs.1989.1.1.1>
- Roberts, Ian. 2010. Varieties of French and the null subject parameter. In Theresa Biberauer, Anders Holmberg, Ian Roberts & Michelle Sheehan (eds.), *Parametric variation: Null subjects in minimalist theory*, 303–327. Cambridge: Cambridge University Press.
- Rooryck, Johan. 1994. On two types of underspecification: Towards a feature theory shared by syntax and phonology. *Probus* 6(2–3). 207–233. <https://doi.org/10.1515/prbs.1994.6.2-3.207>
- Ross, John Robert. 1967. Constraints on variables in syntax. Doctoral dissertation, MIT.
- Safir, Ken. 1986. Relative clauses in a theory of binding and levels. *Linguistic Inquiry* 17(4). 663–689.
- Sportiche, Dominique. 1981. Bounding nodes in French. *The Linguistic Review* 1(2). 219–246. <https://doi.org/10.1515/tlir.1981.1.2.219>
- Spyropoulos, Vasillios & Nikolaos Stamatogiannis. 2011. *Subextraction from subjects in Greek: Things that you can thing and you can say*. Paper presented at the workshop Islands in Contemporary Linguistic Theory, University of the Basque Country, Vitoria-Gasteiz, 16–18 November 2011. http://ehutb.ehu.es/uploads/material/Video/1752/Handout-Spyropoulos____Stamatogiannis.pdf. (26 June, 2018).
- Stepanov, Arthur. 2007. The end of CED? Minimalism and extraction domains. *Syntax* 10(1). 80–126. <https://doi.org/10.1111/j.1467-9612.2007.00094.x>
- Svenonius, Peter. 2004. On the edge. In David Adger, Cécile de Cat & George Tsoulas (eds.), *Peripheries: Syntactic edges and their effects*, 261–287. Dordrecht: Kluwer. https://doi.org/10.1007/1-4020-1910-6_11
- Szabolcsi, Anna. 2006. Strong and weak islands. In Martin Everaert & Henk van Riemsdijk (eds.), *The Blackwell companion to syntax*, Volume 4, 479–532. Malden, MA: Blackwell. <https://doi.org/10.1002/9780470996591.ch64>
- Szabolcsi, Anna & Marcel den Dikken. 2003. Islands. In Lisa Lai Shen Cheng & Rint Sybesma (eds.), *The second Glot international state-of-the-article book: The latest in linguistics*, 213–240. Berlin: de Gruyter.
- Tellier, Christine. 1990. Subjacency and subject condition violations in French. *Linguistic Inquiry* 21. 306–311.
- Tellier, Christine. 1991. *Licensing theory and French parasitic gaps*. Dordrecht: Kluwer. <https://doi.org/10.1007/978-94-011-3596-2>
- Uriagereka, Juan. 1988. On government. Doctoral dissertation, University of Connecticut.

Appendix

This appendix shows speaker types in relation to the preference of CI or ESQ in *wh*-extractions from subject DPs, in descending order according to deviance. The values for CI, ESQ and deviation refer to the Likert scale from 1 to 10 used in the inquiry.

Group 1. No preference (deviance < 0.3)

Speaker	Location	Age/Sex	CI	ESQ	Deviation
12	Lyon (Auvergne-Rhône-Alpes)	23 / f	4.714	5.0	0.286
13	Auvergne (Auvergne-Rhône-Alpes)	26 / f	2.857	2.571	0.286
3	Nice (Provence-Alpes-Côte d'Azur)	26 / m	6.429	6.714	0.285
20	France	25 / m	4.571	4.286	0.285
4	Alençon (Normandie)	28 / m	3.0	2.857	0.143
14	Île de France	27 / f	1.571	1.429	0.142
5	Urzy (Bourgogne-Franche-Comté)	25 / f	2.0	2.143	0.143
2	Val D'Oise (Île de France) / Paris	23 / m	4.429	4.286	0.143
8	Toulouse (Occitanie)	52 / f	1.714	1.714	0
1	Paris	50 / f	1.0	1.0	0
	<i>Average</i>	<i>30.5</i>	<i>3,229</i>	<i>3,20</i>	<i>0.171</i>

Group 2. Preference of CI over ESQ (deviance \geq 0.3)

Speaker	Location	Age/Sex	CI	ESQ	Deviation
16	Besançon (Bourgogne-Franche-Comté)	26 / m	8.714	2.714	6.0
19	Cognac (Nouvelle-Aquitaine)	30 / f	4.714	3.143	1.571
22	Île de France	24 / m	8.286	7.143	1.143
21	Champagnole (Bourgogne-Franche-Comté)	27 / f	3.0	1.857	1.143
11	Reims (Grand Est)	68 / m	4.571	3.571	1.0
15	France	31 / f	2.286	1.714	0.572
10	Beyrouth (Lebanon)	60 / f	2.857	2.286	0.571
17	Montauban (Occitanie)	32 / m	4.286	3.857	0.429
	<i>Average</i>	<i>37,3</i>	<i>4.84</i>	<i>3.286</i>	<i>1.554</i>

Group 3. Preference of ESQ over CI (deviance \geq 0.3)

Speaker	Location	Age/Sex	CI	ESQ	Deviation
18	Amines (Picardie, Hauts-de-France)	28 / f	3.286	6.429	3.143
6	Quebec (Canada)	49 / m	2.571	3.857	1.286
7	Brussels (Belgium)	70 / m	4.571	5.571	1.0
9	Paris	62 / f	5.429	5.857	0.428
	<i>Average</i>	<i>52,25</i>	<i>3.964</i>	<i>5.429</i>	<i>1.464</i>

Moving towards an event

The Romanian prepositional supine construction

Elena Soare

Université Paris VIII / Université Paris Lumières & CNRS

This chapter addresses the status of the prepositional supine construction with the locative preposition *la* “at”, especially as complement of motion verbs.

This construction encodes a goal of motion event, and it interestingly admits a continuation in which this event goal is not reached. In Romanian, locative prepositions have the following properties: (i) they select bare nouns; (ii) they can introduce dative arguments; (iii) in accomplishments, they undo the *telos* of the predicate when introducing an argument. The analysis proposed in this contribution provides a unified account for these three properties, in which the preposition selecting a bare noun introduces a non-achieved goal meaning in the shape of a [-bounded] Path.

Keywords: Romanian, supine, locative prepositions, irresultative predicates, dative, conative construction

1. Introduction

Like Latin, Romanian has a nominal-verbal construction traditionally called ‘supine’. This construction appears, among others, as goal of motion, with the locative preposition *la* “at”, which replaces the case marking on the Latin supine cf. (1a)–(b).¹

- | | | |
|--------|--------------------------|------------|
| (1) a. | <i>Abiit piscatum.</i> | (Latin) |
| | he.went fishing | |
| | “He went fishing.” | |
| b. | <i>Merge la pescuit.</i> | (Romanian) |
| | he.goes at fishing | |
| | “He is going fishing.” | |

1. There is no direct inheritance of the Latin supine in Romanian. For details, see Dragomirescu (2013).

Locative (and other) prepositions are followed by bare nouns in Romanian (Mardale 2008). *La* “at” also introduces dative arguments, and, in accomplishments, it has the property of undoing the *telos* of the main predicate.² The question is how to unify these three properties while accounting for the status of *la* “at” in the constructions shown in (1) and (2). While (1) shows the supine goal of motion construction, (2a) illustrates the conative construction, (2b) the dative, and (2c) a supine expressing a location. In all these constructions it is possible to have a non-culminative continuation; in this respect, the three examples in (2) involve irresultative predicates.

- (2) a. *A citit la roman (dar nu l-a terminat).*
 he.has read at novel (but not it-has finished)
 “(S)he read at the novel (but didn’t finish it).”
- b. *Am dat de mâncare la copii (dar n-au mâncat).*
 I.have given of food to children but not-they-have eaten
 “I gave food to the children (but they didn’t eat).”
- c. *Sunt la pescuit (dar nu am început să pescuiesc).*
 I.am at fishing but not have started to fish
 “I’m out for fishing (but didn’t start to fish).”

On the basis of data as in (1) and (2), I argue that:

- i. the supine in prepositional adjuncts is not verbal but a bare noun.
- ii. the lexical preposition with the bare noun contributes the goal of motion meaning and introduces the idea that the (event) goal is not reached.
- iii. the properties of *la*-supines and contexts in (2a)–(b) can all fall under a unified analysis.

2. Types of supine constructions

2.1 Clausal and prepositional supine

A first thing to observe is that the participial form occurring in Romanian sentences like (1b) is preceded by the preposition *la* “at”; in other constructions, it can take various other prepositions (like *pentru* “for”, *după* “after”). Being preceded by prepositions, correlated with the absence of a determiner, is a property of what ‘verbal supine’ stands for in traditional grammars. Even the new Academy Grammar (Pană Dindelegan 2008; Guțu Romalo 2008, vol. 1: 619) distinguishes between a nominal supine, a nominalization taking the definite determiner, and the verbal supine preceded by prepositions, as in (3a)–(b) respectively (our examples):

2. A *telos* (from the Greek τέλος for “end”, “purpose”, or “goal”) is an end or purpose/intention.

- (3) a. *Fumatul este periculos pentru sănătate.*
 smoking.the is dangerous for health
 “Smoking is dangerous for health.”
- b. *Trebuie să mă las de fumat.*
 I.have to me quit de smoking
 “I have to quit smoking.”

This traditional perspective is maintained in recent grammars, including Pană Dindelegan (2013), where the preposition, even when it is subcategorized by the main verb, is considered to be a “subordinator of non-finite forms” (p. 207).

However, there are two different contexts which are not distinguished by traditional grammars: the supine may be built with a genuine, lexical preposition, or with a functional particle delimiting a clausal domain. To distinguish between the two, note that extraction is possible out of supine clausal complements but not out of prepositional supines. In (4a), the supine is headed by a functional particle (for concreteness, we may assume it to have the status of C° , as proposed in Soare 2002), and correspondingly allows extraction (4b). In (5), the supine is headed by a lexical preposition *pentru* “for” and thus does not accept extraction as shown by (5b).

- (4) a. *E greu de scris articole fără calculator.*
 it.is difficult to write articles without computer
 “It is difficult to write articles without a computer”
- b. *Ce e greu de scris fără calculator?*
 what is difficult to write without computer
 “What is difficult to write without a computer?”
- (5) a. *Calculatorul e pentru scris articole.*
 computer.the is for write articles
 “The computer is to write articles.”
- b. **Ce e calculatorul pentru scris?*
 what is computer.the for write
 (Intended) “What is the computer for writing?”

The supine clausal complement appears in the following contexts: reduced relatives (6a); complex verbal constructions (6b); *tough*-constructions (6c). In all these contexts, the supine is preceded by a *C-de*, which cannot alternate with anything else:

- (6) a. *exemple de reținut*
 examples de remember
 “examples to remember”
- b. *Am de citit.*
 I.have de read
 “I have to read.”

- c. *Romanul este greu de citit.*
 novel.the is tough to read
 “the novel is tough to read”

On the other hand, the prepositional supine is preceded by lexical prepositions, subcategorized by the main predicate:

- (7) a. *Am plecat la pescuit.*
 I.have gone at fishing
 “I’m out for fishing.”
 b. *undiță pentru pescuit*
 rod for fishing
 “(a) fishing rod”

The contrast in (8)–(9) shows that the preposition is selected by the main verb in (9), but not in (8); in (9) but not in (8), the supine alternates with a P-NP construction. *A avea* “to have” and *a termina* “to finish” select for the clausal supine but do not accept a PP; *a se apuca* “to start” selects a prepositional supine or a PP with a nominal.

- (8) a. *Am de citit vs *de carte.*
 I.have de reading de book
 “I have *de read* vs *de book*.”
 b. *A terminat de citit vs (*de) carte(a).*
 has finished de reading (*de) book(.the)
 “(S)he has finished *reading* vs *(*de) (the) book*.”
 (9) *S-a apucat de citit vs de carte.*
 se-he.has taken de reading de book
 “(S)he has started *reading* vs *the book*.”

It is clear, therefore, that we have to distinguish a CP-supine from a PP-supine: they show different properties with respect to selection and extraction.

2.2 Verbal and nominal supine

There is an empirical generalization for Romanian stating that in prepositional contexts and by default, the nominal complements have to be bare by default (Mardale 2008), unless the noun is modified.

This is true for locative prepositions:

- (10) a. *Am plecat la școală / la spital vs *la școala, *la spitalul.*
 I.have gone to school to hospital at school.the at hospital.the
 “I am going to *(*the) school* / to *(*the) hospital*.”

- b. *Școala este lângă /aproape de spital* vs **spitalul*.
 school.the is next-to near of hospital hospital.the
 “The school is next to / near (*the) hospital.”
- c. *Casa este pe deal* vs **pe dealul*.
 house.the is on hill on hill.the
 “The house is on (*the) hill.”

However, there is an exception, namely the preposition *cu* (“with”): in this case the determiner is obligatory:

- (11) a. *Am rămas acasă cu mama* (vs **cu mamă*).
 I.have stayed home with mother.the with mother
- b. *Am subliniat cifrele cu stiloul* (vs **cu stilou*).
 I.have highlighted numbers.the with pen.the with pen
- c. *Am plecat la plimbare cu vaporul* (vs **cu vapor*).
 I.have gone at strolling with boat.the with boat

The same situation arises in the case of bare prepositional supine constructions, Ps obeying the same selectional restrictions as regular nouns:

- (12) a. *M-am apucat de citit (cartea)*.
 me-I.have started of reading book.the
 “I have started reading (the book).”
- b. *Sunt ocupat cu cititul (cărții)*.
 I.am busy with reading.the book.the.GEN
 “I’m busy reading (the book).”

Here is a list of verbs that select a prepositional construction:

- (13) *a se apuca (de)* “to start”, *a se ține (de)* “to keep doing something”, *a se lăsa (de)* “to stop, quit”, *a merge (la)* “to go to”

Verbs that select a functional *de*-supine construction are:

- (14) *a avea* “to have”, *a fi* “to be”, *a termina* “to finish”

For more arguments in favor of this classification, see Soare (2002).

In combination with verbs such as those listed in (14) above, the supine enters a restructuring construction through complex-predicate formation and amounts to a truncated clause with no subject position. The upper layers (e.g., the tense projection and the subject position) are contributed by the first verb which restructures with the truncated supine clause, in turn responsible for lexical aspect and the introduction of the internal argument. The supine with a functional *de* is also present in *tough*-constructions and reduced relatives, where I assume it is also a truncated clause.

From an external-distributional point of view then, we must assume a tripartite classification of supine constructions: (i) definite supine nominal (15); (ii) ‘prepositional’ bare supine nominal (16); (iii) verbal supine (17).

- (15) *Fumatul trabucurilor i-a ruinat sănătatea.*
 smoking.the cigars.GEN him-it.has ruined health.the
 “Smoking cigars ruined his health.”
- (16) *S-a lăsat de fumat (trabucuri).*
 se-he.has left of smoking cigars
 “He has quit smoking (cigars).”
- (17) *Are de citit douăzeci de cărți.*
 he.has of reading twenty of books
 “(S)he has to read twenty books.”

The verbal supine in class 3 constructions allows clitics to be hosted by the first verb, which the bare nominal supine in class 2 does not accept: this is the reason for the ungrammaticality of (18b) below. This proves that the two constructions are fundamentally different. I assume that the supine in the verbal class (iii) constructions is a truncated clause involving complex predicate formation with the first verb, as the clitic climbing suggests.

- (18) a. *L-am terminat/ avut de cules.*
 it-I.have finished had to harvest-SUP
 “I finished/had to harvest/ing it.”
- b. **L-am renunțat la cules.*
 it-I.have give-up at harvest-SUP
 (Intended) “I gave up harvesting it.”

Given the strict parallelism between regular prepositional phrases with nouns and prepositional phrases with the supine, and the differences between (i) contexts with prepositional supines and (ii) contexts with complex predicate formation, I conclude that the supine in the prepositional construction is a bare eventive noun. In this view, *la cules mere* “at picking apples” and *la cules de mere* “at picking of apples” are both nominal but to different degrees. In the first case, we have a bare NP with an incorporated object, while in the second case we have a classical NP modified by a PP. They both have to be distinguished from verbal contexts with complex predicate formation like *am de cules mere* “I have to pick apples”. This is important for the analysis to be proposed, in which the irresultative interpretive component results from the combination between the preposition introducing the nominal supine, and the fact that the supine is a bare noun.

2.3 Definite and bare supine

In work by Alexiadou et al. (2010), among others, the nominal supine is considered to be inflected for imperfective aspect and introduce pluractionality in the context of the definite determiner. I assume that the structure of the prepositional supine is that of a bare eventive noun, i.e., more reduced than the one of the definite supine.

In the definite supine nominal, which has been the object of detailed scrutiny in Iordăchioaia & Soare (2009, 2011, 2015), the definite determiner meets an outer Aspect projection, resulting in a pluractional meaning. This is visible in (19) and (20) respectively by the fact that the supine involves distributivity effects with plurals and in the case of unbounded predicates like states it requires a bounding function, in order to further apply the pluractional operator. The semantic plurality of events introduced by the supine through the contribution of a pluractional operator located in an AspP projection induces ungrammaticality with a singular object in the case of one-time events like *kill* in (19):

- (19) *ucisul* **unui* *journalist/ jurnaliștilor*
 killing.the a.GEN *journalist* *journalists.GEN*
 “killing a *journalist/journalists*”

Moreover, with stative predicates (which are unbounded) the supine is ungrammatical. However, when bounded by a bounding function ‘until’, it becomes grammatical and denotes a habit. These facts diagnose pluractionality. In support of this analysis, one can also note that the definite supine always shifts the aspectual value of the verbal basis into a plurality of events. For more details, see Iordăchioaia & Soare (2009, 2011, 2015).

- (20) a. **statul* *lui Ion la Maria*
 staying.the of *Ion* at *Mary*
 “John’s staying at *Mary’s*”
 b. *statul* *lui Ion la Maria până dimineața târziu*
 staying.the of *Ion* at *Mary* until morning late
 “John’s staying at *Mary’s* until late in the morning”

Unlike the definite supine nominal, the bare supine nominal does not force the pluractional reading, which, when present, is contributed by the main verb. So, (21a) has an episodic one-event reading, while (21b) has a habitual reading, showing that the aspectual value is determined by the first verb (inchoative with *a se apuca* “to begin” and habitual with *a se ține* “to keep ...-ing”) and not by the supine.

- (21) a. *Abia s-a apucat de mâncat carnea.*
 hardly se-he.has taken of eating meet.the
 “(S)he hardly started to eat the meat.”
- b. *Se ține de vânat rațe.*
 se he.keeps of hunting ducks
 “(S)he keeps hunting ducks.”

I thus conclude that unlike the definite supine nominal, the bare supine nominal only presents inner-aspectual (atelic) specifications. There is no reason to think that the bare supine nominal presents an AspP layer; the fact that the presence of adverbs and prepositional aspectual adjuncts is questionable seems to indicate that such a projection is absent in the bare prepositional supine. In (22a), the PP *în cinci minute* “in five minutes” cannot be interpreted as modifying the supine but only the main verb. As an indication, we can note that it is only possible to question the main verb and not the supine, as shown by (22b):

- (22) a. *# S-a apucat de mâncat carnea în cinci minute.*
 se-he.has started of eating meat.the in five minutes
- b. *Când s-a apucat de mâncat?* vs *# Când a mâncat?*
 when se-he.has started of eating when he.has eaten
 “When did (s)he start to eat?” vs “When did (s)he eat?”

As an interim conclusion, I propose that the prepositional supine construction involves a bare nominal supine with lexical-aspectual specifications (it is atelic). Both atelicity and the absence of the determiner are important factors in the make-up of the goal-of-motion construction, in addition to the semantics of the lexical preposition which introduces the supine.

I now turn to a comparison between the prepositional supine construction and the other two constructions mentioned in the Introduction, namely the conative construction and the prepositional dative.

3. The contribution of the lexical preposition

3.1 The conative construction, the defeasible telos and the prepositional dative

Certain lexical (locative) prepositions with bare nouns yield defeasible accomplishments (Bar-el et al. 2005, among others), more particularly in what is commonly called the conative construction (Levin 1993):

- (23) a. *A citit romanul (*dar nu l-a terminat).*
 he.has read novel.the but not it-he.has finished
 “(S)he read the novel (*but didn’t finish it).”
- b. *A citit la roman (dar nu l-a terminat).*
 he.has read at novel but not it-he.has finished
 “(S)he read at the novel (but didn’t finish it).”

These lexical prepositions have a decisive contribution undoing the *telos* of the main predicate. The question is what exactly they contribute in a goal-of-motion construction and in the defeasible accomplishment.

Note that the same preposition is used to mark dative in colloquial Romanian, with the possible continuation implying that the Theme did not reach the Goal. On the one hand, there is a contrast between inflectional dative in (24b) and the prepositional *la*-dative in (24a) with respect to the irresultative continuation, which, according to my consultants, is only possible with the prepositional dative (24b). On the other hand, when the verb entails that the Theme reaches the Goal, like in the case of *înmâna* “to hand”, the *la*-dative is ungrammatical (25c). Moreover, as (25d) shows, with this kind of verb the irresultative continuation is not possible.

- (24) a. *Am dat de mâncare la copii (dar n-au mâncat).*
 I.have given of food at children but not-they.have eaten
 “I gave food to the children, but they did not eat.”
- b. *Am trimis scrisoarea la asociație (dar n-au primit-o).*
 I.have sent letter.the to association but not-they.have receive it
 “I sent the letter to the association, but they didn’t receive it.”
- (25) a. *Am dat de mâncare copiilor (?dar n-au mâncat).*
 I.have given of food children.DAT but not-they.have eaten
 “I gave food to the children, ?but they did not eat.”
- b. *Am trimis scrisoarea asociației (?dar n-au primit-o).*
 I.have sent letter.the association.DAT but not-they.have receive it
 “I sent the letter to the association, but they didn’t receive it.”
- c. **Am înmânat diploma la elevi.*
 I.have handed diploma.the to pupils
- d. *Am înmânat diploma elevilor (*dar n-au primit-o).*
 I.have handed diploma.the pupils.DAT but not-they.have received-it
 “I handed in the diploma to the students, *but they didn’t receive it.”

What we therefore observe is an alternation between a structure with a full DP in an object position involving culmination, and a P+bare NP construction in which the *telos* is not reached. The question is how this difference is achieved.

3.2 The prepositional supine and the unachieved goal

The non-culmination continuation appears in a large number of prepositional constructions, introducing Goal bare nouns, including the prepositional supine.

- (26) a. *Sunt la fumat.*
I.am at smoking
“I’m out for smoking.”
- b. *Mașina e pentru condus invitații la gară.*
car.the is for take guests.the to station
“The car is to take the guests to the station.”
- c. *Mașina e pentru plimbare.*
car.the is for going-out
“The car is for going out in.”
- (27) a. *Sunt la fumat, dar încă nu am început să fumez.*
I.am at smoking but yet not I.have started to smoke
“I’m out for smoking but didn’t start smoking yet.”
- b. *Mașina e pentru condus invitații la gara dar încă nu
i-am condus.*
them-I.have driven
“The car is to drive guests to the station but we never driven them yet.”
- c. *Mașina e pentru plimbare, dar nu ne-am plimbat încă cu ea.*
car-the is for strolling but not us-we.have strolled yet with it
“The car is to go out but we didn’t go out with it yet.”

There is a contrast with the definite supine, which does not admit the non-culmination continuation:

- (28) **Am început cititul, dar încă nu citesc.*
“I have started reading but am not reading yet.”

Note again that all these constructions involve article drop. I suggest that the preposition replaces case marking, which in Romanian generally depends on the determiner. There is alternation between a genuine case (dative, for instance) and the Prep+bare NP construction. I assume certain locative Prepositions select for an NP – the construction is not always definite contra Mardale (2008) and Dobrovie-Sorin & Giurgea (2013). The examples below show that prepositions like *pe* “on” or *la* “at” select bare NPs, which cannot be referred back with demonstratives (29), unlike prepositions like *lângă* “near” or *spre* “towards”, which select covert definite DPs, which can be referred back with demonstratives.

- (29) a. *Casa e pe deal. *Acesta este lângă o pădure.*
 house.the is on hill that is near a forest.
- b. *Am plecat la pescuit. *Acesta/*aceasta e o activitate foarte plăcută.*
 I.have gone at fishing this/that is an activity very pleasant
- (30) a. *M-am așezat lângă profesor/ (Acesta) era beat.*
 me-I.have sat near professor this.one was drunk
- b. *Se îndreaptă spre școală/ (Aceasta) e deschisă.*
 se he.heads toward school this.one is opened

The fact that the preposition combines with a bare NP is important in this discussion. We can see that the conative construction involves alternation between a definite DP in an argument position which is an affected Theme, and a PP selecting a non-definite NP which cannot denote an affected Theme and thus does not involve telicity. Therefore, the ‘unachieved’ meaning component must be traceable back to this definite / non-definite alternation and to the contribution of the lexical preposition.

3.3 Unfolding the unachieved goal meaning

The alternation between a definite DP object and a PP+bare noun brings about the distinction between the achieved and the unachieved goal meaning. A definite DP is merged in an argument position of the main predicate (for instance, in terms of Borer 2005, in Spec,AspQ), yielding the telic interpretation. A PP+bare noun is not inserted in this position, but probably in an adjunct position, and does not yield a telic interpretation. The unachieved goal meaning involves a couple of ingredients, more precisely the meaning of the preposition (“at”) and article drop. Alternatively, in Ramchand’s (2008) terms, the PPs would be Rhemes which denote unbounded Paths, while definite DPs introduce bounded Paths. Thus, what I would like to propose here can be spelled out in different syntactico-semantic approaches.

I thus claim that a goal-of-motion component is contributed to the main predicate by the lexical preposition in the shape of a [–bounded] PathP (31), signifying that the (event) goal denoted by the bare noun is not achieved. This happens for the goal-of-motion bare supine examples above, and is also true in the case of *pentru* “for” with ordinary nouns and with bare supines.

- (31)
- | | |
|-----------------|--------------------|
| PathP | |
| Path [–bounded] | NP |
| la/pentru | plimbare / pescuit |
| at/for | walk / fishing |

Inside the PathP, the [-bounded] locative Preposition marks case on the bare NP. As we saw, the absence of the article is an important ingredient in the non-achieved goal meaning. I assume that the article is completely missing and is not incorporated into the Preposition (the nominal is not a covert definite cf. above). Article drop in the supine amounts to its truncated structure. D selects AspP in the pluractional supine but is absent here, and the supine does not project AspP, but only encodes lexical atelic aspect. This in turn is also an important component of the unachieved goal meaning (a-*telos*).

Combining this PathP construction with a verbal predicate amounts to the absence of a result state in the construal. In studies about spatial expressions, (un)boundedness is an important ingredient (cf. Jackendoff 1985; Hale & Keyser 2002). Spatial adpositions may denote a bounded or unbounded path or trajectory. In other terms, notions of central and terminal coincidence (Hale 1986) have been used to distinguish telic and atelic predicates; central coincidence is identified with atelic predicates, while terminal coincidence is identified with bounded path predicates. *La* + bare NP in Romanian introduces exactly the former type of meaning in the constructions at interest here.

The analysis proposed here can be extended to the prepositional colloquial dative and the conative construction in Romanian, which involve the same ingredients: a [-bounded] PathP with a bare (atelic) noun which is inserted in an adjunct position and not in the argument position. This amounts to the non-existence of a result state, and yields the unachieved goal meaning. If this is correct, the pattern *la* “at” + bare NP would have a unified contribution to the verbal predicate, which one could label the unachieved goal construal.

4. Conclusion

In this chapter, I addressed the parallelism between three types of constructions involving the preposition *la* “at” in Romanian: the conative construction, the colloquial dative construction and the supine goal-of-motion construction, which all admit a non-culminative continuation. I have first shown that the supine in the goal-of-motion construction is a bare noun of events. The goal-of-motion construction involves locative prepositions with bare nouns, which amounts to undoing the *telos* of the embedded predicate. This construction comes with a [-bounded] Path component inducing the (unachieved) goal of motion and thus the result state is not projected any more. Article drop participates in the non-achieved goal meaning, the non-definite status of the nominal contributing to the non-affectedness in the

case of a Theme (conative construction) or to the unachieved goal meaning in the case of a Goal of motion construction or a Benefactive (dative alternation). Hence, the analysis can be extended to the prepositional colloquial dative and the conative construction in Romanian.

References

- Alexiadou, Artemis, Gianina Iordăchioaia & Elena Soare. 2010. Number/aspect interactions in the syntax of nominalizations: A distributed morphology approach. *Journal of Linguistics* 46. 537–574. <https://doi.org/10.1017/S002226710000058>
- Bar-el, Leora, Henry Davis & Lisa Matthewson. 2005. Non-culminating accomplishments. *Proceedings of NELS* 35. 79–92.
- Borer, Hagit. 2005. *Structuring sense, Volume 2, The normal course of events*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199263929.001.0001>
- Dobrovie-Sorin, Carmen & Ion Giurgea (eds.). 2013. *A reference grammar of Romanian*, Volume 1, The noun phrase. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.207>
- Dragomirescu, Adina. 2013. *Particularități sintactice ale limbii române în context romanic. Supinul*. Bucharest: Editura Muzeului Național al Literaturii Române.
- Guțu Romalo, Valeria (ed.). 2008. *Gramatica limbii române*, 2 volumes. Bucharest: Editura Academiei Române.
- Hale, Kenneth. 1986. Notes on world view and semantic categories: Some Warlpiri examples. In Pieter Muysken & Henk van Riemsdijk (eds.), *Features and projections*, Volume 25, 233–254. Dordrecht: Foris. <https://doi.org/10.1515/9783110871661-009>
- Hale, Kenneth & Samuel J. Keyser. 2002. *Prolegomenon to a theory of argument structure*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/5634.001.0001>
- Iordăchioaia, Gianina & Elena Soare. 2009. Structural patterns for plural blocking in Romance nominalizations. In Enoch Aboh, Elisabeth van der Linden, Josep Quer & Petra Sleeman (eds.), *Romance languages and linguistic theory: Selected papers from 'Going Romance' 2007*, 145–160. Amsterdam: John Benjamins. <https://doi.org/10.1075/rlt.1.08ior>
- Iordăchioaia, Gianina & Elena Soare. 2011. A further insight into the syntax-semantics of pluractionality. *Proceedings of SALT*, 95–114. <https://doi.org/10.3765/salt.v21i0.2625>
- Iordăchioaia, Gianina & Elena Soare. 2015. Pluractionality with lexically cumulative verbs: The supine nominalization in Romanian. *Natural Language Semantics* 23(4). 307–352. <https://doi.org/10.1007/s11050-015-9117-9>
- Jackendoff, Ray. 1985. *Semantics and cognition*. Cambridge, MA: MIT Press.
- Levin, Beth. 1993. *English verb classes and alternations. A preliminary investigation*. Chicago: University of Chicago Press.
- Mardale, Alexandru. 2008. Prépositions et article défini en roumain. In Adèle de Saint-Pierre & Mélanie Thibeault (eds.), *Actes des 21èmes Journées de Linguistique*, 78–93. Québec: Centre interdisciplinaire de recherche sur les activités langagières, Universitaire de Laval.
- Pană Dindelegan, Gabriela. 2008. Supinul. In Valeria Guțu Romalo (ed.), *Gramatica limbii române*, Volume 1, 509–524. Bucharest: Editura Academiei Române.

- Pănă Dindelegan, Gabriela (ed.). 2013. *The grammar of Romanian*. Oxford University Press.
- Ramchand, Gillian Catriona. 2008. *Verb meaning and the lexicon. A first phase syntax*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511486319>
- Soare, Elena. 2002. *Le supin roumain et la théorie des catégories mixtes*. Doctoral dissertation, Université de Paris VIII.

Cyclicity without containment in Romanian perfects

Donca Steriade

Massachusetts Institute of Technology

The Romanian perfect exhibits a form of directional paradigm uniformity: verbal perfect forms adopt the stress and segmental characteristics of the perfect participle. An analysis of this pattern of paradigmatic identity is proposed, which has broader implications for the theory of the phonological cycle.

Keywords: cycle, paradigm uniformity, base-derivative correspondence

1. Introduction

This study analyzes a pattern of similarity among Romanian verb forms, previewed in (1).

(1) Stem identities in the Romanian perfect¹

	PERFECT			NON-PERFECT	
	3sg. perf.	1PL perf.	Participle	3sg. indic.pres.; gerund	
a. “fall”	[kʌzú]	[kʌzú]-ram	[kʌzút]	[kád]-e	[kʌz]-índ
“burn”	[árs]-e	[árs]-e-ram	[árs]	[árd]-e	[arz]-índ
b. “hold”	[tsinú]	[tsinú]-ram	[tsinút]	[tsín]-e	[tsin]-índ
“put”	[pús]-e	[pús]-e-ram	[pús]	[pún]-e	[pun]-índ

1. Data sources for this study include Lombard & Gâdei (1981); Pană Dindelegan (ed.) (2013); and dextonline.ro, an extensive, searchable lexical database. The IPA symbols used here are mapped to Romanian graphs as follows: [ʌ, ɨ, ʃ, ts] = <ă, î, ș, ț>, [tʃe, tʃi] = <ce, ci >, [ke, ki] = <che, chi>.

The left side of the table in (1) shows four sets of perfect forms with identical stress and segmentally similar stems. The perfect participles display an unpredictable difference between an *-s* suffix, as in *ars*, and a *-t* suffix, as in *kazút*. Some properties of participial stems are transmitted to the tensed perfect: witness *pús*, *púse* vs *tsinút*, *tsinú*. The non-perfect forms, shown on the right side of (1), have stress and segmentals that differ from the perfect. In non-perfect verb forms, stress alternates and generally abides by the constraints applicable to morphologically simple forms, as argued below. The perfect differs.

I will show that the stress and segmental composition of all perfect forms can be predicted from the perfect participle, which follows the accentual pattern of simple words. This requirement of stem identity that governs the entire paradigm of the Romanian perfect is analyzed here as an instance of cyclic inheritance, in which the participle functions as the Base (= cycle *n*) and tensed perfect verbs are generated as its Derivatives (= cycle *n*+1). Unlike in standard cyclic cases, the Base in these perfect paradigms is not contained, morphologically or syntactically, in its Derivatives. A modified theory of the cycle (Stanton & Steriade manuscript) does justice to this and comparable other cases.

2. Romanian perfects

Romanian perfect paradigms consist of a participle (PPf) and three sets of verbal forms (VPf): a simple perfect comparable in its aspectual value to the French *passé simple*, a pluperfect and an analytic perfect, comparable to the *passé composé*, consisting of the PPf plus an auxiliary. Of interest here are the synthetic verb forms, the simple perfect and pluperfect.

The table in (2) presents two complete paradigms, accompanied by a morphological parse of the tensed perfects, one for each of the types shown. 'PERF₁₋₃' is a reference to different perfect exponents; *se* is the pluperfect suffix, *ra* is a marker of plurality in perfects; AGR refers to all other person-number endings.²

2. The function of *-e-* in sigmatic perfects is to block impossible C clusters, as in *ars-e-se*, and to promote anti-homophony, as in 3rd SG *ars-e*, which would otherwise merge with *ars*, the PPf.

(2) Identities in the stems of perfect forms of two strong verbs

Infinitive	<i>kad-eá</i> ROOT-THEMEV “fall”		<i>árd-e</i> ROOT-THEMEV “burn”		
PPf	<i>kaz-ú-t</i> ROOT-PERF ₁ -PERF ₂		<i>ár-s</i> ROOT -PERF ₃		
VPf:	1	[kaz-ú]-j	[kaz-ú]-ra-m	[ár-s]-e-j ³	[ár-s]-e-ra-m
simple	2	[kaz-ú]-fj	[kaz-ú]-ra-tsj	[ár-s]-e-fj ³	[ár-s]-e-ra-tsj
perfect	3	[kaz-ú]	[kaz-ú]-ra	[ár-s]-e	[ár-s]-e-ra
VPf:	1	[kaz-ú]-se-m	[kaz-ú]-se-ra-m	[ar-s]-é-se-m	[ar-s]-é-se-ra-m
pluperfect	2	[kaz-ú]-se-fj	[kaz-ú]-se-ra-tsj	[ar-s]-é-se-fj	[ar-s]-é-se-ra-tsj
	3	[kaz-ú]-se	[kaz-ú]-se-ra	[ar-s]-é-se	[ar-s]-é-se-ra

ROOT-PERF₁-(se_{PLUPERF})-(r_{APL})-AGR ROOT-PERF₃-e-(se_{PLUPERF})-(r_{APL})-AGR

Like (1), the data in (2) shows that stress is on the same stem syllable in VPfs as in the participle. (2) also illustrates an exception to the general pattern of accentual identity in the perfect: stress changes in verbal forms to avoid identical *sese* strings. Thus, pluperfects like *arsésem* contain *sése*, not **ársesem*, the expected form given *árseram*. This limited deviation from identity between the PPf and the VPf forms will play a role in the analysis. (1) and (2) also show that perfect stems are segmentally identical, aside from the PPf suffix *-t*, in forms like *kaz-ú-t*. This *-t* is systematically missing in VPf forms for reasons explored below.

The verbs in (1–2) are two of the ca 250 strong verbs originating in the 2nd and 3rd Latin conjugations. The vast majority of Romanian verbs descend from the Latin 1st and 4th conjugations, in *-ā*, *-ī*. As in Latin, the perfects of these verbs, two of which appear in (3), preserve the theme vowel of the present. By contrast, the strong verbs lose their present theme vowels in the perfect.

(3) Perfect forms of two weak verbs: “hear” and “praise”

Infinitive	<i>auz-í</i> ROOT-THEMEV “hear”		<i>lud-á</i> ROOT-THEMEV “praise”		
PPf	<i>auz-í-t</i> ROOT-THEMEV -PERF ₂		<i>lud-á-t</i> ROOT -THEMEV-PERF ₂		
VPf: simple	1	[auz-í]-j	[auz-í]-ra-m	[lud-á]-j	[lud-á]-ra-m
perfect	2	[auz-í]-fj	[auz-í]-ra-tsj	[lud-á]-fj	[lud-á]-ra-tsj
	3	[auz-í]	[auz-í]-ra	[lud-á]	[lud-á]-ra

Weak verbs like those in (3) follow in the perfect the same patterns of identity as those seen in (1–2). The unique exception from identity is lawful. It appears in the

3. For this paradigm type, the 1st and 2nd sg forms admit a variant with final stress: *arséj*, *árséj*. This is the only option reported by Zafiu (2013: 33). The variant with root stress, *árséj*, *árséj*, is reported elsewhere (Lombard & Gâdei 1981: 135) and has been encountered by the present writer numerous times, including in rimes like *ramásej* (remain-Vpf-1SG) - *kásej* (house-DAT SG), where the unreduced vowels and the requirements of riming identity guarantee the location of stress.

3rd SG.PF. of verbs like *lud-á*, which are realized with a change of theme vocalism, [*lud-á*]→ [*lud-á*], to avoid homophony with the imperfect 3rd SG. [*lud-á*].

The stem identity seen in (1–3) is limited to the perfect. Non-perfect stems differ segmentally from each other and from the perfect. Their stress alternates:

(4) Alternations in non-perfect forms: “fall” and “hear”

pres.ind.	1	<i>kád</i>	<i>kɒd-é-m</i>	<i>a.úd</i>	<i>a.uz-í-m</i>
	2	<i>káz-j</i>	<i>kɒd-é-tsj</i>	<i>a.úz-j</i>	<i>a.uz-í-tsj</i>
	3	<i>kád-e</i>	<i>kád</i>	<i>a.úd-e</i>	<i>a.úd</i>

I show next that the accentual mobility in the non-VPf forms is the effect of rankings holding generally in the language. What will have to be explained is the contrast between the accentual invariance characteristic of the perfect and the regular accentual mobility observed outside the perfect.

3. The stress system outside the perfect

Outside the perfect, stress is largely predictable in Romanian. Most words are stressed on the penult, unless the final is heavy, in which case final stress is the rule: see constraints and rankings in (5a–c). Three additional options are attested, but disfavored:⁴ (i) antepenult stress in words with light finals and penults, like *kámera* “room”; (ii) penult stress when the final is heavy, e.g., *úmar* “shoulder”; and, even less commonly, (iii) final stress on a light syllable, e.g., *halvá* “halvah”. I analyze all these deviations from the general pattern by letting a lexically indexed constraint, IDENTSTRESS_{LEX} (Pater 2000) outrank some of the M constraints, as in (5e). Pre-antepenult stress and antepenult stress in words with closed penults or finals are virtually impossible in native words and nativized loans. I use the conjunction of WSP and *LAPSER to analyze this, (5g–h). I use grid-based constraints (Gordon 2002) and constraint conjunction, in (5g), but neither is critical to the main argument.

- (5) Stress constraints and rankings for mono-morphemes
- NONFINALITY (NF): one * for any final stress.
 - WEIGHT-TO-STRESS (WSP): one * for every stressless heavy syllable.
 - WSP >> NONFINALITY >> STRESSRIGHT
 - *LAPSER: one * for each final pair of stressless syllables, 00#
 - IDENTSTRESS_{LEX} >> *LAPSER – WSP# >> NONFINAL >> IDENT STRESS

4. Vasiliu (1965) provides lexical counts on stress in roots, from which stress in fully inflected words can be inferred. See also Steriade (1985); Chițoran (2001).

- f. *EXTLAPSER: one * for 000#.
 g. *LAPSER-WSP one * any string that violates both LAPSER and WSP.
 h. *EXTLAPSER, *LAPSER-WSP >> IDENT-STRESS_{LEX} >> WSP, *LAPSER

In non-perfect forms, the rankings proposed above cause accentual alternations, as in *mút- λ* , *véd-e* “displaces/sees” vs *mut-ám*, *ved-ém* “we change/we see”.

(6) a. *mút- λ* “changes”

Root: <i>mut-</i>	NONFIN	STRESSR
<i>mutá</i>	*!	
☞ <i>múλ</i>		*

b. *mut-ám* “we change”

Root: <i>mut-</i>	WSP	NONFIN
☞ <i>mutám</i>		*!
<i>múλam</i>	*!	

In verb roots with lexical stress on a non-final root syllable, *LAPSER-WSP causes alternations between antepenult stress, in V-final forms like *mátur- λ* , and final stress in words ending in VCC₀, like *matur-ám*, *matur-índ*.

(7) a. *mátur- λ* “he/she sweeps”

<i>mátur-</i>	IDSTRESS _{LEX}	*LAPSER
<i>matúra</i>	*!	
☞ <i>mátura</i>		*

b. *matur-ám* “we sweep”

<i>mátur-</i>	*LAPSER-WSP	IDSTRESS _{LEX}
<i>máturam</i>	*!	
<i>matúram</i>		**!
☞ <i>maturám</i>		*

The distribution of stresses in Romanian is identical to that of Spanish, as analyzed by Harris (1983). Our analyses differ in an interesting way. Harris uses segment extrametricality to generate two of the three marked stress patterns of Spanish-Romanian. In his analysis, antepenult stress results from a final vowel being extrametrical; penult stress in words with heavy finals is due to an extrametrical consonant. When applied to Romanian, Harris’s proposal will parse *kámera* as *káme.r-< λ >* – with angle brackets marking extraprosodicity – and *úmar* as *úma<r>*. When nothing is extrametrical, the default pattern emerges. The lexical exceptions to stress in verbs bear on the difference between analyses. Verb roots are frequently followed by monosyllabic endings. On our analysis, any syllable in the root can be lexically stressed. This marked stress is protected, modulo higher constraints, by IDENTSTRESS_{LEX}. This idea is reflected in the rankings in (5h). It generates all and only the patterns of exceptions found in single forms.

Consider now the attested patterns of accentual alternations. Those found in verbs are seen in (8). They include an unmarked alternation pattern, (8a), in which stress moves between a final VCC₀ and a penult followed by a light final, plus two marked types, in (8b–c). (8b) illustrates verbs with marked antepenult stress; (8c) displays verbs with fixed penult stress. Finally, row (d) illustrates a conceivable but

impossible pattern, to be compared with the attested ones: antepenult stress before a light final, alternating with penult stress before a heavy final.⁵

(8) Normal and exceptional stress alternations in present tense verbs

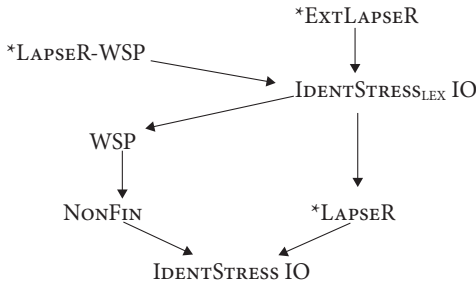
	Stress distribution	Ø ending	_V ending	-VC ending	
a.	default: alternating penult/final stress	<i>arát</i>	<i>arát-a</i>	<i>arát-ám</i>	“show”
b.	marked: 1st root σ / final stress	<i>mátur</i>	<i>mátur-a</i>	<i>mátur-ám</i>	“sweep”
c.	marked: fixed stress on 2nd root σ	<i>adúk</i>	<i>adútf-e</i>	<i>adútf-em</i>	“bring”
d.	unattested: 1st root σ /2nd root σ	* <i>bátur</i>	* <i>bátur-a</i>	* <i>batúr-am</i>	

The analysis proposed in (5) generates all the attested patterns of accentual alternations, (8a–c), and only those. Patterns (8a) and (8b) were derived in (6) and (7); pattern (c) follows from the same rankings, if roots like *adúk*, *adútf-em* contain a lexically stressed second syllable. The fact that (8d) is impossible has been previewed in (7), and follows from the assumption that any accentually irregular root has *some* underlying stress. If the lexical stress is on the second root syllable, we expect (8c); if on the first, we expect (8b). The shift in (8d) cannot be generated from one lexically stressed syllable, in first or second position; nor from multiple lexical stresses.

To understand how Harris’s analysis might generate (8), recall that his proposal is to derive exceptional stress by marking a word-final segment extrametrical. In verbs, this word-final segment may belong to an ending, but the extrametricality condition must be entered in the lexical entry of roots, to distinguish type (a) roots from type (b) or (c). Setting this odd feature aside, a Harris-style analysis can’t generate patterns (b) or (c), because neither of them displays a constant pattern of segment extrametricality: e.g., *mátur*, *mátura* in (8b) can be generated by extrametricality of the last segment, but that also predicts **matúram*. Similarly, *adúk* and *adútf-e* get the correct stress if nothing is extrametrical, but *adútfem* requires that final /m/ be extrametrical. The only case that is easily derived by Harris’s analysis is the unattested pattern (d): **bátu.<r>*, **bátu.r<a>*, **batúra<m>*. We conclude from this that the better analysis of marked and default stress is based on the ranking schema $M_1 \gg \text{IDENT}_{\text{STRESS}_{\text{LEXIO}}} \gg M_2$. That analysis is summarized below:

5. Pre-antepenult stress is made impossible, in our analysis, by undominated *EXTLAPSER.

(9) Ranking summary



4. Perfect correspondence: Accentual evidence

Let us return to the pattern of accentual similarity between perfect stems. The analysis in (9) helps identify a key difference between VPfs and PPfs. The verbal forms systematically violate active Markedness constraints: forms like [kaʒ-ú] violate NONFINALITY; pluperfects like [kaʒ-ú]-se-ra, [kaʒ-ú]-se-ra-m, violate *LAPSER and *LAPSER-WSP. PPfs, by contrast, are regular in terms of (9): all end in a heavy syllable and all are stressed on that heavy final. From this difference between the accentually unmarked participles and the accentually marked tensed perfects, we infer that the perfect similarity pattern is directional: stress in the participle is computed according to (9); it is then transferred to the VPfs, which deviate from (9) in order to preserve their accentual similarity to the PPf.

A simple mechanism generates this form of directional identity: the PPf is the Base (B) of the entire perfect paradigm. Its shape is generated in a first step, comparable to a cycle 1. VPf forms are its Derivatives (D): they use the surface form of the PPf, as the input into their evaluation. The constraint requiring accentual identity among correspondent perfect forms, IDENT STRESS (BD), ranks above all Markedness constraints in (9).

(10) a. Deriving the Base, PPf: *kaʒ-ú-t* “fallen”

UR: /kad-u-t/	WSP	NONFIN
<i>kázut</i>	*!	
☞ <i>kazút</i>		*

- b. Deriving Ds: *kazú* “fell3SG”, *kazú-se-ra-m* “fell-plupf-1PL”

B: [káz-ú-t] “fallen”	IDENT STRESS (BD)	NONFIN
[kázú]	*!*	
☞ [kázú]		*

B: [káz-ú-t] “fallen”	IDENT STRESS (BD)	*LAPSER-WSP
[kázú]-se-rá-m	*!	
☞ [kázú]-se-ra-m		*

The contrast between *matur-á-m*, (7b), with stress shifted to the final, and *kazú-se-ra-m*, (10b), with stem stress, emerges from the common ranking schema characterizing one OT theory of cyclic phonology (Benua 1997): FAITH (BD) >> M >> FAITH (IO). In the present case, the schema is instantiated as: IDENT STRESS (BD) >> *LAPSER-WSP >> IDENT STRESS_{LEX} (IO).

A deviation from accentual identity in the perfect was mentioned earlier, in the case of pluperfects in *sese*, e.g., Ppf [ár-s] vs [ar-s]-és-e “had burned”. Here stress is shifted off its expected position in the root. The trigger of this shift is an identity avoidance constraint: a change of stress is used to differentiate two otherwise identical syllables, to reduce their similarity. I lack evidence for the full scope of this constraint in Romanian. The version in (11a) prohibits only identical *CeCe* strings. More plausibly, any sequence of strictly identical CVs is avoided.⁶

- (11) a. OCP (*CeCe*): one * for each identical *CeCe* string.
 b. OCP (*CeCe*) >> IDENT STRESS (BD)
 c. Deriving pluperfect *ars-é-se* “had burnt-3SG”

B: [árs] ‘burnt’	OCP (<i>CeCe</i>)	IDENT STRESS (BD)	*LAPSER
[árs]-e-se	*!		*
☞ [ars]-é-se		* (á-a)	

For *sese*-pluperfects other than the 3rd SG, this analysis is still insufficient. The stress shift to the first *se* is predicted by the present analysis in the 3rd SG because the second *se* is a final light. But other *sese*-forms have an incentive to shift stress differently, because their Markedness score improves if stress lands on the second *se*. (12) shows how the attested 1st PL form [ars]-é-se-ra-m loses to penult-stressed *[ars]-e-sé-ra-m. The 3rd pl. [ars]-é-se-ra will similarly lose to *[ars]-e-sé-ra, due to a critical *LAPSER violation.

6. Zukoff (2015) motivates an identical constraint in Ponapean. For Romanian, searches in dexoline.ro for identical *CeCe* strings show that they are absent word internally, outside the onomatopoeic and child-directed lexica. Items like *zile-le* “the days” are clitic groups.

- (12) An initial failed attempt to derive
- ars-é-se-ra-m*
- “burnt-plupf-1PL”

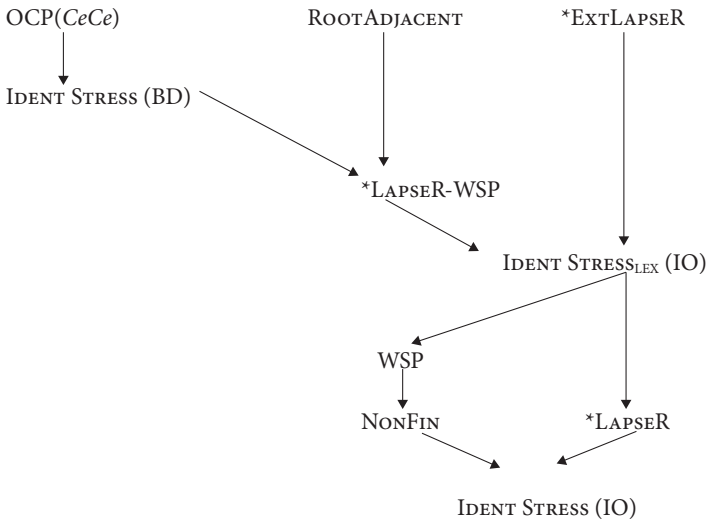
B: [árs] “burnt”	OCP (CeCe)	ID STRESS BD	*LAPSER-WSP
[árs]-e-se-ra-m	*!		*
[ars]-e-se-rá-m	*!	* (á-a)	
[ars]-é-se-ra-m		* (á-a)	*!
! ➔ [ars]-e-sé-ra-m		* (á-a)	

The simplest remedy is to stipulate that stress must remain in proximity of the root, not separated from it by any syllable. A constraint implementing this idea is **ROOT ADJACENT**, which bans candidates whose stress is separated by a syllable or more from the root. **ROOT ADJACENT** eliminates in (13) the winner of (12), allowing the attested form to emerge as optimal. A summary ranking follows.

- (13) Deriving pluperfect
- ars-é-se-ra-m*
- “had burnt-1PL”

B: [ars]	OCP (CeCe)	ROOTADJ	ID STRESS (BD)	*LAPSER-WSP
☞ [ars]-é-se-ra-m			*	*
[ars]-e-sé-ra-m		*!	*	

- (14) Perfect correspondence



5. Alternatives

The present analysis explains an asymmetry noted earlier: the surface structure of the PPF transparently explains its stress – e.g., *kaɫz-ú-t* has predictable stress on a heavy final – but the identical stress of corresponding VPfs lacks surface phonotactic justification. Thus [*kaɫz-ú*] “fell-3SG” surfaces instead of the expected *[*káɫz-u*], as does [*kaɫz-ú*]-*ram*, instead of expected *[*kaɫz-u*]-*rám*, like *matur-ám*. The preceding section has laid out the beginnings of an account that explains this asymmetry: the participle is the base of the VPfs, the tensed perfect forms, so the latter get their stress from the former, and they pay a Markedness price – e.g., violations of *LAPSER-WSP – for their identity to their Base. I explore next two alternative accounts of this identity, one based on McCarthy’s (2005) Optimal Paradigms (OP) theory, and one based on the idea of lexically stressed morphs.

We begin with the latter. It is conceivable that unexpected stresses like [*kaɫz-ú*] “fell-3SG”, [*auz-í*] “heard-3SG” are due to the presence of lexically stressed theme vowels, *ú* and *í*. If so, the stress in [*kaɫz-ú-t*] “fallen”, [*auz-í-t*] “heard” is doubly motivated, both by lexical stress and by the weight of the final, while stress in [*kaɫz-ú*], [*auz-í*] is due entirely to the theme vowel’s underlying accent, and not to B-D correspondence. We should examine then the evidence for lexical stress on theme vowels. The theme vowel *-ú-* of [*kaɫz-ú*] is limited to the perfect, where no independent consideration suggests that it has, or lacks, inherent stress. More revealing are the weak verbs of the type *auz-í*, *lɫud-á*, *arɫt-á*, seen earlier in (3) and (8). These have the same theme vowels in the perfect and the present. Their perfect paradigms were seen in (3). Representative present forms follow.

(15) Present of weak verbs

Infinitive	<i>auz-í</i> ROOT-THEMEV “hear”	<i>arɫt-á</i> ROOT-THEMEV “show”
present	1 [aúð]	[<i>auz-í</i>]- <i>m</i> [<i>arɫt</i>] [<i>arɫt-á</i>]- <i>m</i>
	2 [aúz]- <i>j</i>	[<i>auz-í</i>]- <i>tsj</i> [<i>arɫts</i>]- <i>j</i> [<i>arɫt-á</i>]- <i>tsj</i>
	3 [aúð]- <i>e</i>	[aúð] [<i>arát</i>]- <i>ɔ</i> [<i>arát</i>]- <i>ɔ</i>

This data shows that, aside from the infinitive and the perfect, the theme vowels are stressed only if they happen to occur in predictably stressed positions: e.g., present [*auz-í*]-*tsj* or [*arɫt-á*]-*tsj*, both with stress on heavy finals. The same theme vowels can surface in stressless form: the [ɔ] in the 3rd persons of the present [*arát*]-*ɔ*, [*láud*-ɔ] is identifiable as the theme vowel *-a-* of the 1st conjugation. In such forms, the theme vowel is regularly unstressed by the system in (14), and is regularly reduced. The final *-e* of 3rd sg. *aúð-e* could similarly be a stressless avatar of the theme vowel *-i-*. The theme vowels delete in the present of the weak conjugations, if followed by vowel-initial endings, some of which then proceed to disappear in

turn: e.g., /aud-i-u/, /arat-a-u/, /arat-a-i/ are the underlying forms of 1SG *aúd* and *arát* and of 2nd SG *arátsj*.⁷ Significantly, stressed vowels do not otherwise delete in Romanian.⁸ All this suggests that the theme vowels of weak verbs are not invariably stressed. Lexical stress is then not a credible basis for an account of the accentual identities in the perfect, quite aside from the fact that it would account for only a fragment of the perfect identities we analyze.⁹

The remaining alternative to our analysis uses the theory of Optimal Paradigms (OP; McCarthy 2005; cf. related proposals in Kenstowicz 1998), according to which a set of non-directional correspondence constraints promote similarity between the stems of forms that comprise a lexeme's inflectional paradigm. The OP constraint set includes MAX (OP), DEP (OP), and IDENT F (OP), for any feature, including stress. Their function is to verify that each pair of stems in an inflectional paradigm is identical for the property named in the constraint. Entire candidate paradigms are evaluated simultaneously. For each constraint C, violations of C incurred by individual members are summed over each candidate paradigm. This includes violations of the OP CORR constraints. Each such violation represents a pair of paradigm members whose stems differ in the relevant respect.

A successful OP alternative to the current analysis could invalidate the key claim of this study, that asymmetrical Base-Derivative correspondence obtains even when the Base is not a constituent in its Derivatives. In OP analyses, there are no Bases and no Derivatives. There are only members of the same paradigm seeking to converge upon the optimal compromise between stem identity across the paradigm and Markedness/IO Faithfulness satisfaction. If such an analysis is right, my claim about B-D correspondence without containment cannot be sustained.

7. An account along these lines is defended by Feldstein (1994). The *-u* ending of the 1st SG is justified in Steriade (1985). Significantly, Feldstein finds evidence for some underlyingly stressed morphemes (the imperfect *-á-*) but not for underlyingly stressed theme vowels. A question that remains open is how to reconstruct the opacity inherent in Feldstein's analysis – V₁-deletion in hiatus followed by deletion of the trigger V₂ – in the present account.

8. See Steriade (1985) on stress-conditioned vowel deletion and gliding in Romanian.

9. The infinitive is also stressed on a final light in most verb types: e.g., *arat-á*, *auz-í*. A possible reason for these final stresses is that these infinitives are truncated from regularly penult-stressed forms like *arat-á-re*, *auz-í-re*, old infinitives which now function as verbal nouns. A truncation account of stress in the current infinitives is defensible synchronically and requires no mention of lexically stressed vowels. An alternative explanation is that final stress in infinitives is a means to avoid homophony to other paradigm cells. For verbs like *arat-á* and *auz-í*, the regular penultimate stress will generate, when we plug in the reduction processes affecting atonic syllables, *arát-A* and *aúz-j*, but these forms are already in use as, respectively, 3rd sing and 2nd sing. indicative presents. Paradigm-internal anti-homophony has significant other effects in the Romanian conjugation. Either way, the upshot is, again, that no evidence supports lexically stressed theme vowels.

In (16), I present a successful OP analysis of the perfect paradigm of “fall”, seen earlier in (2). Three candidate paradigms are worthy of consideration. The paradigm in (16a) displays regular stress on heavy finals, and otherwise on penults, but suffers disqualifying violations of OP IDENT STRESS, a constraint defined in (17). The paradigm in (16b) has uniform initial stress, but too many of its forms suffer from lapse. Finally, (16c) is accentually uniform, like (16b), and reduces Markedness violations to a minimum. It wins.

(16) OP analysis of stress in a perfect paradigms: *kazút* “fallen”

	/kazu-/	OP IDSTRESS	*LAPSER	NF
a.	<i>kazút;</i> <i>kazúj, kazúfj, kázu</i> <i>kazurám, kazurátsj, kazúra;</i> <i>kazusém, kazusefj kazúse,</i> <i>kazuserám, kazuserátsj, kazuséra</i>	98*!		9*
b.	<i>kázut,</i> <i>kázuj, kázuj kázufj,</i> <i>kázurəm, kázurətsj, kázura;</i> <i>kázusem, kázusefj kázuse,</i> <i>kázuserəm, kázuserətsj, kázuserə</i>		9*!	
☞ c.	<i>kazút;</i> <i>kazúj, kazúfj kazú,</i> <i>kazúrəm, kazúrətsj, kazúra;</i> <i>kazúsem, kazúsefj kazúse,</i> <i>kazúserəm, kazúserətsj, kazúserə</i>		3*	4*

(17) OP IDENT STRESS: In each pair of paradigm-internal correspondent forms, W_1 - W_2 , assign a * for each V in W_1 that has a correspondent V' in W_2 such that V and V' differ in stress.

The test of this type of analysis comes when some constraint promoting a failure of identity among paradigm members dominates OP CORR. In the present case, one dominant constraint is OCP(*CeCe*), which blocks strictly identical *sese* sequences. To satisfy OCP (*CeCe*) and ROOTADJACENT, the first *se* must be stressed. (18) reveals that, when the dissimilarity-inducing constraint OCP (*CeCe*) outranks OP IDENT, the optimal candidate is one that minimizes the numbers of pairwise dissimilarities by shifting as many stresses as possible to the post-root syllable. The resulting forms are all accentually identical, with just one item, monosyllabic PPF *árs*, inevitably left with root stress. This winning candidate is very different from the actual paradigm of such perfects, represented by (18c): real VPfs deviate from

identity to the PPF only when forced by OCP(*CeCe*). This suggests that the number of violations of OP IDENT is in fact irrelevant. Of the other candidates in (18), (a) represents the paradigm with regular stress, on heavy finals and otherwise on penults; (b) represents the candidate that fully satisfies OP IDENT STRESS, by keeping stress on the root syllable. Each of them is eliminated by the top two constraints of the analysis. The significant part, though, is that the actual winner, (18c) is also eliminated, by OP IDENT.

(18) Failed OP analysis of stress in a perfect paradigms: *árs* ‘burnt’

	/ars-/	OCP (<i>CeCe</i>)	OP IDENT STRESS	*LAPSER	NF
a.	<i>árs;</i> <i>ars-éi, arsé-fj, árse</i> <i>arse-rám, arse-rátsj, arsé-ra;</i> <i>arse-sém, arse-séffj arsé-se,</i> <i>arse-serám, arse-serátsj, arse-séra</i>	2*!	44*		9*
b.	<i>árs;</i> <i>árs-ej, árse-fj, árse</i> <i>árs-e-rám, árse-rátsj, árse-ra;</i> <i>árs-e-sem, árse-seffj, árse-se,</i> <i>árs-e-serám, árse-serátsj, árse-séra</i>	6*!		12*	*
c.	<i>árs;</i> <i>árs-ej, árse-fj, árse</i> <i>árs-e-rám, árse-rátsj, árse-ra;</i> <i>arsé-sem, arsé-seffj arsé-se,</i> <i>arsé-serám, arsé-serátsj, arsé-séra</i>		84*!	6*	*
!☛d.	<i>árs;</i> <i>ars-éj, arsé-fj, arsé</i> <i>arsé-rám, arsé-rátsj, arsé-ra;</i> <i>arsé-sem, arsé-seffj arsé-se,</i> <i>arsé-serám, arsé-serátsj, arsé-séra</i>		24*	3*	3*

I noted in fn. 3 the existence of accentual variation in the stress of perfects like *árse*: variants like *ars-éj*, *arsé-fj* exist for the 1st and 2nd sg, alongside *árs-ej*, *árs-efj*. They can be generated using additional constraints that outrank IDENT STRESS (BD). Importantly, such variants do not support an OP analysis either. Whatever constraint determines their stress, its interaction with OP IDENT will predict in the OP analysis the wrong form **arsé*. The right generalization is that the stress of each VPF form remains identical to that of its PPF, independently of what the rest

of the paradigm does. Constraints like OCP (*CeCe*), which shift stress away from the Base have a local effect on individual forms, not a global effect on the paradigm. This pairwise relation between the participle and individual VPf forms can only be derived in a directional B-D analysis.

This discussion explains why the second alternative to our B-D analysis is not viable. What it does not invalidate is an OP analysis of fragments of the perfect paradigm. If we consider just the tensed perfect, or just the pluperfect, in isolation from other perfect forms, the accentual uniformity within these microparadigms is complete, and thus analyzable in OP. But what is striking about the Romanian perfect is the nearly complete accentual – and, as we shall see, segmental – identity of perfect forms *across* the PPf-VPf divide and across the perfect-pluperfect boundary. OP does not help generate that.

6. Containment and the Romanian perfect

Thus far I have shown that directional B-D Correspondence is the right mechanism to generate the pattern of identity described here. The next step in the overall argument is to show that the Base of the perfect paradigm, the PPf, is not contained in its VPf Derivatives.

In a surface phonological sense, this is a directly observable fact. In the paradigm of PPf *kazút*, the VPfs are *kazú*, *kazúram*, *kazúseram*, etc., not **kazút*, **kazútram*, **kazútseram*. Let us restate the significance of this fact for the analysis. The perfect marker *-t* of *kazút* causes its *u* to be stressed, by making heavy the final syllable. Then this *-t* is indirectly responsible, via IDENT STRESS (BD), for the stresses of VPfs like *kazú*, *kazúram*, *kazúseram*, etc: without a requirement of identity to regularly stressed *kazút*, the stresses of these VPf forms would be elsewhere. Since *-t* itself is missing in VPfs, this shows that *kazút*, the full form whose stress is cyclically transmitted to the VPf forms, is not phonologically contained in them.

But we had set out to establish a distinct point, which is more directly relevant to the broader conclusion of this study: the Base is not *syntactically* embedded in its Derivatives in these perfect paradigms. This section completes that argument. I show that *-t* cannot occur in VPfs like *kazú*, *kazúram* because a syntactic feature expressed by *-t* is not compatible with the syntactic structure of VPf forms. If any syntactic feature of the PPf is not contained in a VPf, it follows that the former is not embedded in the latter. Once we establish this, a non-containment-based account of two central facts will have to be provided: PPfs and VPfs are in correspondence; and, in this relation, the PPfs are the Bases.

6.1 Classes of strong perfects and perfect exponents

There are six varieties of PPfs, three from weak verbs and three from strong verbs. In such forms, the suffixes *-t*, *-s* and *-u* (in 19a–b) are the perfect markers whose distribution we investigate, while *-a-* and *-i/i-* (in 19c) are aspectually neutral theme vowels used in the present and perfect of weak verbs. The numbers in (19a) are counts of strong verbs from the lists in Lombard & Gàdei (1981).

(19) a. Classes of strong perfects and lexical counts

	Class 1 (N: 32)	Class 2 (N: 186)	Class 3 (N: 21)
Infin.	<i>kad-eá</i> < <i>kad-é</i> “fall”	<i>árd-e</i> “burn”	<i>fiérb-e</i> “boil”
PPf	<i>kaz-ú-t</i>	<i>ár-s</i>	<i>fiér-t</i>
VPf	[<i>kaz-ú</i>]- <i>j</i> , [<i>kaz-ú</i>]- <i>ra-m</i> , ...	[<i>ár-s</i>]- <i>e-j</i> , [<i>ár-s</i>]- <i>e-ra-m</i> , ...	[<i>fiér-s</i>]- <i>e-j</i> , [<i>fiér-s</i>]- <i>e-ra-m</i> , ...

b. Distribution of perfect markers in the strong verbs

	PPf			VPf		
	Class 1	Class 2	Class 3	Class 1	Class 2	Class 3
-t	√(<i>kaz-ú-t</i>)		√(<i>fiér-t</i>)			
-u	√(<i>kaz-ú-t</i>)			√(<i>kaz-ú</i>)		
-s		√(<i>ár-s</i>)			√(<i>ár-s-e</i>)	√(<i>fiér-s-e</i>)

c. Classes of weak perfects and examples

	-a verbs	-i verbs	-i verbs
Infin.	<i>lud-á</i> “praise”	<i>auz-í</i> “hear”	<i>hotar-í</i> “decide”
PPf	<i>lud-á-t</i>	<i>auz-í-t</i>	<i>hotar-í-t</i>
VPf	[<i>lud-á</i>]- <i>j</i> , [<i>lud-á</i>]- <i>ra-m</i> , ...	[<i>auz-í</i>]- <i>j</i> , [<i>auz-í</i>]- <i>ra-m</i> , ...	[<i>hotar-í</i>]- <i>j</i> , [<i>hotar-í</i>]- <i>ra-m</i> , ...

The suffix distribution in (19) suggests several elements of analysis. First, the suffix *-t* never occurs in VPf forms, whether weak or strong, only in PPfs. We explain this fact if *-t* is an exponent of both [+perfect] and of a lexical category feature, say [+adjective], which is incompatible with *-t*'s occurrence in VPfs.

Second, the suffix *-s* occurs in both VPfs and in PPfs, as seen in the strong classes 2 and 3, (19a–b). We explain this if *-s* is an exponent of [+perfect] only, and thus useable in both adjectival and verbal forms. Similarly, the suffix *-u* occurs in both VPf forms and PPf forms, as seen in the strong class 1. We infer that it too is an exponent of [+perfect] only, like *-s*.

Third, the suffixes *-u* and *-s* do not co-occur: there are no **kaz-ú-s* type perfects. We have explained this by attributing to *-u* and *-s* identical exponence functions. There is no need for both in one form.

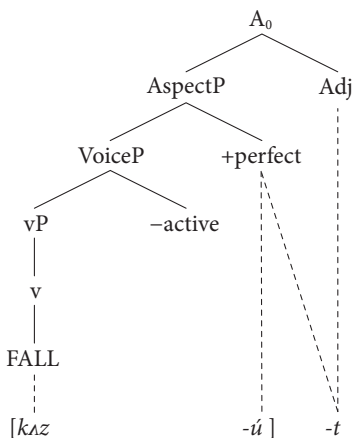
Fourth, the suffixes *-u* and *-t* do co-occur in PPfs, as in *kaz-ú-t*. Here, *-t* would be a sufficient exponent, but *-u*'s function is to satisfy segmental phonotactics and to keep stress off the root, a point developed in § 5.3.

Next, the suffixes *-s* and *-t* do not co-occur either. This is partly explained by the syntactic features our analysis attributes to them. In VPfs, *-t* is unusable, qua adjectival suffix. For PPfs, *-s* expresses a proper subset of the features expressed by *-t*. By the Subset Principle (Halle 1997), we expect then only *-t* to occur in PPfs, and certainly not a combination of *-t* and *-s*. What remains unexplained is what makes *-s* a possible PPf marker at all, for some strong verbs: one might expect *-t* to always replace it, as in class 3 perfects. The analysis proposed here offers no synchronic answer to this question. All sigmatic PPfs are inherited archaisms, rather than innovations. Synchronically, they must be analyzed by means of lexically indexed exponence constraints.

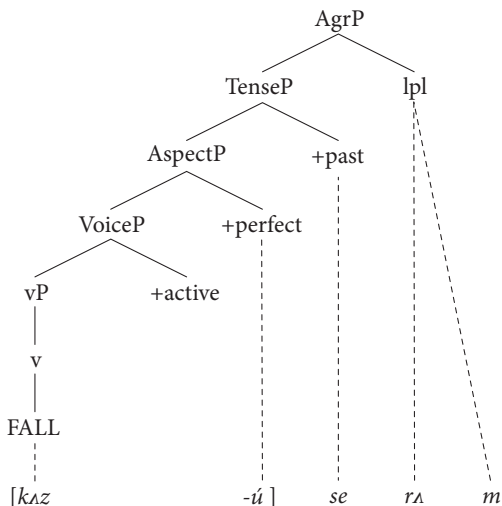
To sum up, this section has proposed that *-t* is an exponent of the features [+perfect] and [+adjective], while *-s* and *-u* express only [+perfect]; verb roots must be lexically indexed in order to use the [+perfect] exponent *-s*. This feature assignment explains several distributional generalizations, including the fact of central interest here: the *-t* of PPfs like *kazút* can't be present in corresponding VPfs like *kazú*, *kazúseram*, while the *-s* of participles like *árs*, must be present in VPfs like *árse*, *árseram*. (20) helps visualize how syntactic structures map to exponents under this proposal. Dotted lines connect each morph to the nodes it is an exponent of. Perfect stems are in brackets.

(20) Exponents of two perfect forms: A PPf and a VPf (1st pl pluperfect)

a.



b.



Our next step is to clarify what causes sets of perfect forms, such as {*kazút; kazú, kazúseram...*}, to enter into correspondence at all. The answer will be the hypothesis that such instances of paradigm uniformity are due to constraints that place in correspondence cognate sets of forms sharing a syntactic feature, or a bundle of syntactic features, independently of syntactic embedding. In the present case, the feature requiring correspondence is [+perfect].

6.2 Perfect correspondence

The constraint causing stem identity among Romanian perfect forms is (21)

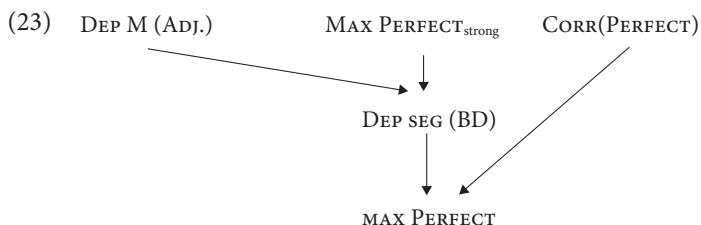
- (21) CORR(PERFECT): If two syntactic structures S_1 and S_2 are lexically related and contain the value [+perfect], the stems of the word-sized exponents of S_1 and S_2 stand in correspondence.

When sets of perfect forms enter in correspondence in order to satisfy (21), constraints like MAX, DEP and IDENT are activated. Under certain rankings, they will generate paradigms with phonologically identical or similar stems. The version of CORR(PERFECT) in (21) requires correspondence between ‘stems’. We define ‘stem’ in this context as the smallest contiguous string of morphs that includes exponents of all the syntactic nodes shared by S_1 and S_2 . How a perfect stem maps to a set of syntactic nodes was illustrated in (20), and reveals that the syntactic nodes corresponding to stems placed in correspondence by (21) need not be identical: the voice features of PPfs and VPfs differ.

CORR(PERFECT) causes not only accentual but also segmental identities between perfect stems. We revisit the data in (19) to illustrate one of these. In class 3 strong perfects – e.g., *fiert, fierse, fiersera* – the adjectival *-t* must be replaced in VPfs by *-s*, because bare VPfs like **fiere, *fierera*, without any perfect suffix, are unacceptable. The table in (19b) reveals a related generalization: every strong VPf contains some overt perfect suffix. This is not the case in weak verbs, as (19c) and (3) show: weak PPfs like *auzit* ‘heard’ correspond to VPf sets that lack any perfect marker, like *auzi* ‘heard-3SG’. Alternative VPfs like **auzú*, using the perfect suffix *-u*, or simple perfects like **auzís(e)*, using *-s* – the latter comparable to attested strong paradigms like {*skris, skrise, skrísera ...*} ‘written, wrote’ – are impossible in the weak verbs. The analysis must characterize two complementary generalizations: it must exclude all strong perfects lacking any [+perfect] affix, while also insuring that weak VPfs lack [+perfect] morphemes. The first step is to introduce a constraint, based on Wolf’s (2008) MAX MORPH schema, banning perfect structures that lack exponents of the perfect.

- (22) MAX PERFECT: a * for every instance of [+perfect] that lacks an overt exponent.

An undominated version of (22) will be indexed to the class of strong verbs in (19a). This accounts for the first generalization: all strong perfects contain an exponent of the perfect. To explain the absence of perfect markers in weak VPfs, i.e., the impossibility of weak paradigms like *{*auzít, auzú, auzúra...*} or *{*auzít, auzise, auzísera...*} (where **auzise*, etc., is a simple perfect), we appeal to CORR(PERFECT). In weak paradigms, the perfect stem of PPfs differs from that of VPfs minimally, only as required by the need to exclude *-t* from tensed forms. In impossible paradigms like *{*auzít, auzú, auzúra*} or *{*auzít, auzise, auzísera*} (the latter two as simple perfects) the stems of participial and verbal forms differ more from each other: *-t* is not just missing from all VPfs, but is also replaced by *-u* or *-s*. To analyze the weak paradigms then, we rank CORR(PERFECT) and DEP SEG (BD) above the general version of MAX PERFECT. The constraint DEP M (ADJ.), also based on Wolf's (2008) proposals, completes the analysis: it bans VPfs containing exponents of [+adjective] like *-t*.



Below I derive individual VPf members, from a weak and a strong perfect paradigm, to illustrate how (23) characterizes some of the generalizations about perfect exponence presented in this section. As before, I assume that the PPf is the Base of each paradigm. Perfect stems are in brackets. The effect of MAX PERFECT_{strong} is shown in (24a), a 3rd PL VPf from a strong verb. The contribution of CORR(PERFECT) is seen in (24b), a 3rd PL VPf from a weak verb: CORR(PERFECT) activates DEP SEG (BD), which excludes two losing candidates. Without CORR(PERFECT), DEP SEG (BD) would not block satisfaction of MAX PERFECT, and forms like *{*auzí-s*-(*e*)-*ra*}, second candidate in (24b), would win.

(24) a. Deriving one VPf member of a strong perfect paradigm, class 3.

B: [<i>fiér-t</i>] 'boiled'	DEP M (ADJ.)	MAX PERFECT _{strong}	DEP SEG (BD)
[<i>fiér-t</i>]- <i>e-ra</i>	*!		
[<i>fiérb</i>]- <i>e-ra</i>		*!	*
[<i>fiér</i>]- <i>e-ra</i>		*!	
☞ [<i>fiér-s</i>]- <i>e-ra</i>			*

b. Deriving a VPf member of a weak perfect paradigm.

B: [auzí-t] ⁱ ‘heard’	DEP M (ADJ.)	CORR(PERF)	DEP SEG (BD)	MAX PERF
[auzí-t] ⁱ -(e)-rλ	*!			
[auzí-s]-(e)-rλ		*!		
[auzí-s] ⁱ -(e)-rλ			*!	
[auz-ú] ⁱ -rλ			*!	
☞ [auzí] ⁱ -rλ				*

6.3 Excursus: Perfect-infinitive correspondence

We have yet to explain the double perfect exponence seen in PPfs like *kaz-ú-t*, *dur-ú-t*, *par-ú-t*, of the strong class 1. Our analysis states that the perfect suffix *-u* expresses a proper subset of the syntactic features expressed by *-t*. We should ask then why *-u* is used at all in the *-u-t* participles, where *-t* makes it redundant: why *kazút*, *parút*, *durút* and not **káz**t*, **dú**r**t*, **pár**t*, like *fiért*?

A further correspondence effect provides an answer. If the infinitive of a verb has iambic stress (as in *kad-eá*, *par-eá*, *dur-eá*), the PPf has iambic stress as well (as in *kaz-út*, *par-út*, *dur-út*); and similarly for weak verbs, e.g., *auz-í*, *auz-í-t*.¹⁰ The structure ROOT-V-t insures final, iambic stress in the PPf, in virtue of (14). Without the *-u* in PPfs like *kaz-ú-t*, stress would be on the root, and would mismatch the stressless root in the infinitive of this verb, *kad-eá*. Conversely, if the infinitive has root stress – because the verb root contains a lexical stress – the PPf generally has root stress too: *fiér**b-em*, *fiér-t* (**fiér**bút*); *ár**d-em*, *ár-s* (**ár**z-út*); *pún-em*, *pú-s*; *dútf-em*, *dús*, etc. Had these PPfs included *-u*, in addition to *-t* or *-s*, stress would shift to the suffix, mismatching the infinitive (*pú**ne*, **púnút*), or else WSP would be violated (*pú**ne*, **púnut*). There are exceptions, mostly in *k*-final roots (*fát**Ńe*, *fakút*, perhaps because *kt*, as in expected **fák**t*, would regularly become *pt*), but the general correlation seems clear.

10. One exception is *adáus* ‘added’, with lexical stress in the PPf but regular final-stressed infinitive *adáu**g-á* (VPfs in Lombard & Gàdei 1981 for *adáu**g-á* follow this old PPf: *adáu**se* ‘added-3rd sg’). All other verbs with lexical stress in the weak conjugations behave as described in the text: e.g., *méstek* ‘I chew’, showing lexical stress on the first syllable, but *méstek-á*. The perfect forms *méstek-át*, *méstek-á*, etc. follow the infinitive. What must be explained in such paradigms is the final stress in infinitives: given the lexically stressed root we expect **méstek-a* or, with reduction, **méstek-a*. The hypothesis outlined in fn. 10, that infinitives are truncated derivatives of verbal nouns in *-re* and preserve their stress, explains these forms. *EXTLAPSER blocks **méstek-a-re*, forcing stress to advance to the penult in the *-re* noun. The infinitive preserves that stress.

A preliminary suggestion for the mechanism behind this correlation is a preference for rhythmic identity between all non-tensed forms of a root, i.e., infinitives and participles. This will explain the double exponence in strong PPfs like *kaz-ú-t*. Comparable chains of correspondence are found in Latin verb paradigms and their non-verbal derivatives (Steriade 2016, manuscript). Romanian appears to have inherited such abstract patterns of rhythmic identity, if not their overt Latin manifestation.

7. Base Priority in the perfect and its source

In Sections 2–3 I have presented an argument that the accentual identity among perfect stems in Romanian is directional: the stress of the PPf follows the general stress rules of the language and determines, via B-D correspondence, the stress of the VPfs. This results in accentual anomalies in the stress of the latter. The preceding section has contributed to the same directional hypothesis. It has proposed that the choice to include the *-u* marker in strong PPfs like *kazút* – or not to include it, as in *árs* or *fiért* – stems from correspondence between the non-tensed forms of a root, the infinitive and the PPf. Once it is made, this choice of *-u* vs no *-u* is transmitted to the VPfs, again via B-D correspondence. The direction is the same in both cases, from PPf to VPfs. The directional B-D analyses of stress identity offered in § 3 have thus received independent justification.

Section 5 has shown that this directional identity cannot be attributed to the Base being contained in the Derivative: PPfs are not embedded in cognate VPfs, in either a syntactic or a morphological sense. Hence, syntactic containment is not the source of the B-D asymmetry studied here.

What *is* then the source of this directional effect? We should distinguish grammar-external factors, which favor the selection of some forms as Bases in certain paradigms, from the grammatical reflex of the asymmetry between Bases and Derivatives. As far as the latter goes, the proposed mechanism can be (25), the statement that certain complex expressions have derivational priority over others.

(25) The exponents of PPfs are generated prior to those of VPfs.

(25) is comparable to the assignment of some derivatives to Level 1 and of others to a later Level. Statements like it are presupposed in Stratal OT analyses of level ordering and cyclicity (Kiparsky 2000), in analyses of directional paradigm uniformity effects elsewhere (Hall & Scott 2007) and in other instances of cyclic effects where Base Priority does not stem from syntactic containment (Steriade manuscript, 1999, 2008, 2016).

It is clear that statements like (25) over-generate if nothing limits the pairs of expressions they can link, but it is not clear that this concern is best addressed by designing grammars in a different way. Albright (2010) has identified relative ‘informativity’ as the property responsible for selecting one Base among a set of paradigm cells. The most informative potential Base B in a paradigm is one that would trigger the smallest number of phonological and morphological neutralizations, if B’s stem were generalized to all cells. More informative Bases are favored, because they allow larger numbers of contrasts to surface across paradigms. Albright shows that this factor selects the right Base in a number of diachronic developments that create uniform paradigms. Relative informativity is a grammar-external factor, because there is no intrinsic connection between some paradigm cell being more informative than others and that cell being designated as the Base of a paradigm, in a statement like (25) or in any other format.

A different grammar-external factor that can differentiate potential Bases is token frequency. The more frequent the cell – or the higher the average token frequency of a cell type across paradigms – the better known its properties will be to learners, and thus the more likely it is that learners will extend its properties to other cells. Token frequency plays only an indirect role in Albright’s model, and only this indirect role is validated by his findings. However, the frequency difference between Romanian PPfs and VPfs is so large at present that it may eliminate the VPfs from the competition for basehood. The simple perfect is in the process of being replaced in the literary language by the analytic perfect, which uses the PPf. The remaining synthetic VPf paradigm is the pluperfect, an uncommon tense-aspect combination. An illustration of the frequency disparity between VPfs and PPfs is provided by comparing Google hits for VPf and PPfs in 10 common roots. Results are in (26). The verbs were chosen so as to avoid homophony between the perfect forms and any other paradigm members. This is why only one verb of the productive *-i* conjugation is included: the 3rd SG PERF (e.g., *auzi*) is identical with the infinitive and is spelled identically with the 2nd SG PRES. As (26) shows, VPf percentages range between a high of 15%, in one verb, and more common figures approaching 0.

- (26) Percentages of VPf forms of 10 verbs from the total number of perfect forms: Google hits in millions, unless indicated otherwise

	<i>încuia</i>	<i>mânca</i>	<i>auzi</i>	<i>încrede</i>	<i>râmânea</i>	<i>desface</i>	<i>cumpăra</i>	<i>înțelege</i>	<i>vedea</i>	<i>cădea</i>
total perfect	778k	96	42	437k	110	2	24	25	28	6
% VPf	0.02	0	0.06	0	0.01	0.03	0.02	0	0.06	0.15

The frequency disparity is even more substantial than (26) suggests. The PPf is phonologically homogeneous, as all its gender-number inflected forms have accentually identical stems, without the benefit of leveling. By contrast, as seen in (16) and

(18), stress in the VPfs will alternate in the absence of some form of paradigmatic leveling, landing on the root, or the aspect markers, or the endings, depending on what affixes follow the root. This enhances the frequency disparity between candidates for the Base stem: the unique stem of the PPf with its cumulative frequency competes against each one of the three or four accentually diverse stems to be expected in the VPfs.

If we set frequency aside, it is unclear that informativity favors the PPf as the base of the perfect paradigm. In most strong verbs, the verb root is directly followed by *-t* or *-s*, participial suffixes that cause extensive neutralizations through the loss of the last consonant, as in *fier* < *fierb-t*, *ars* < *ard-s*, *kurs* < *kurg-s*, *pus* < *pun-s*. Establishing the relative informativity of the PPf compared to other potential bases would require a difficult comparison to a hypothetical set of VPfs, whose stems would have to be computed independently of the participle and of each other. In the absence of a clear way to carry this out, it seems safe to assume that the token frequency disparity is a large enough factor in favor of the PPf to dwarf any possible VPf advantage in informativity.

To summarize, the proposal is to model the B-D asymmetry between participles and tensed perfect forms by the derivational statement in (25). According to (25), VPfs are generated in a second derivational step, post PPf. The constraint CORR (PERF) and the identity conditions it triggers are necessarily activated only at this second stage, and can thus affect only the shape of VPfs. The preliminary proposal is to let extra-grammatical factors, like comparative token frequencies of various candidate Bases, dictate the derivational order in (25) and comparable cases.

The most important point in this discussion is that an analysis like the one advocated here is equally available to standard cases of cyclic inheritance, where Bases *are* contained in their Derivatives. The difference between those patterns and the one studied here (or in Hall & Scott 2007; or in the diachronic developments reconstructed by Albright 2005, 2010) is that (a) derivational priority statements like (25) are unnecessary in standard cyclic cases, but required here; and (b) a preference can be detected in other instances for B-D correspondence to involve only Bases that are contained in their Derivatives. The Romanian perfect case shows, along with much other evidence, that such a preference for B-D correspondence under containment is violable: any constraint penalizing non-nested correspondent pairs must rank below CORR(PERFECT) in Romanian. Once we recognize the violability of this containment condition, B-D correspondence can receive an identical analysis as paradigmatic uniformity. In the Romanian perfect case, we have shown that this unification is not just conceptually satisfying, but empirically necessary: § 4 has shown that non-directional mechanisms fail to generate the uniform paradigms described here.

Going beyond the data analyzed in this chapter, Romanian morphology is a rich source of phenomena involving cyclic inheritance without containment. Some were studied elsewhere (Steriade 2008; Stanton & Steriade manuscript), while others await formal analysis. Taken together, they suggest that the constraint limiting BD correspondence to nested pairs of Bases and Derivatives ranks very low throughout the grammar of Romanian.

References

- Albright, Adam. 2005. The morphological basis of paradigm leveling. In Laura Downing, T. Alan Hall & Renate Raffelsiefen (eds.), *Paradigms in phonological theory*, 170–210. Oxford: Oxford University Press.
- Albright, Adam. 2010. Base-driven leveling in Yiddish verb paradigms. *Natural Language & Linguistic Theory* 28(3). 475–537. <https://doi.org/10.1007/s11049-010-9107-z>
- Benua, Laura. 1997. Transderivational identity: Phonological relations between words. Doctoral dissertation, University of Massachusetts Amherst.
- Chițoran, Ioana. 2001. *The phonology of Romanian: A constraint-based approach*. Berlin: Mouton de Gruyter.
- Feldstein, Ronald. 1994–1995. Accentul verbului în limba română. *Dacoromania* 1(1–2). 75–96.
- Gordon, Matt. 2002. A factorial typology of quantity insensitive stress. *Natural Language & Linguistic Theory* 20(3). 491–552. <https://doi.org/10.1023/A:1015810531699>
- Hall, Tracy A. & John H. Scott. 2007. Inflectional paradigms have a base: Evidence from s-dis-similation in southern German dialects. *Morphology* 17. 151–178. <https://doi.org/10.1007/s11525-007-9112-z>
- Halle, Morris. 1997. Distributed morphology: Impoverishment and fission. In Benjamin Bruening, Yoonjung Kang & Martha McGinnis (eds.), *MITWPL 30: Papers at the Interface*, 425–449. Cambridge, MA: MIT Press.
- Harris, J. W. 1983. *Syllable structure and stress in Spanish: A non-linear analysis*. Cambridge, MA: MIT Press.
- Kenstowicz, Michael. 1998. Base identity and uniform exponence. In Jacques Durand & Bernard Laks (eds.), *Current trends in phonology: Models and methods*, 363–394. Salford: University of Salford.
- Kiparsky, Paul. 2000. Opacity and cyclicity. *The Linguistic Review* 17.351–367. <https://doi.org/10.1515/tlir.2000.17.2-4.351>
- Lombard, Alf & Constantin Gâdei. 1981. *Dictionnaire morphologique de la langue roumaine*. Bucharest: Editura Academiei & C.W.K Gleerup.
- McCarthy, John. 2005. Optimal paradigms. In Laura Downing, T. Alan Hall & Renate Raffelsiefen (eds.), *Paradigms in phonological theory*, 170–210. Oxford: Oxford University Press.
- Pană Dindelegan, Gabriela (ed.). 2013. *The grammar of Romanian*. Oxford: Oxford University Press.
- Pater, Joe. 2000. Nonuniformity in English stress: The role of ranked and lexically specific constraints. *Phonology* 17(2). 237–274. <https://doi.org/10.1017/S0952675700003900>
- Stanton, Juliet & Donca Steriade. Manuscript. *Stress and the cycle, in English and elsewhere*. MIT, 2018.

- Steriade, Donca. 1985. Glides and vowels in Romanian. In Claudia Brugman & Monica Macaulay (eds.), *Proceedings of the 10th Annual Meeting of the Berkeley Linguistics Society*, 47–64. Berkeley, CA: Berkeley Linguistics Society.
- Steriade, Donca. 1999. Lexical conservatism in French adjectival liaison. In Barbara Bullock, Marc Authier & Lisa Reed (eds.), *Proceedings of the 25th meeting of the Linguistic Symposium on Romance Languages*, 243–271. Amsterdam: John Benjamins.
<https://doi.org/10.1075/cilt.185.18ste>
- Steriade, Donca. 2008. A pseudo-cyclic effect in Romanian morphophonology. In Asaf Bachrach & Andrew Nevins (eds.), *Inflectional identity*, 314–359. Oxford: Oxford University Press.
- Steriade, Donca. Manuscript. *The cycle without containment: Latin perfects*. MIT, 2013.
- Steriade, Donca. 2016. The morpheme vs. similarity-based syncretism: Latin t-stem derivatives. In Ana Luís & Ricardo Bermúdez-Otero (eds.), *The morpheme debate*, 112–174. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780198702108.003.0006>
- Vasiliu, Emanuel. 1965. *Fonologia limbii române*. Bucharest: Editura Științifică.
- Wolf, Matthew. 2008. Optimal interleaving: Serial phonology-morphology interaction in a constraint-based model. Doctoral dissertation, University of Massachusetts Amherst.
- Zafiu, Rodica. 2013. Mood, tense and aspect. In Gabriela Pană Dindelegan (ed.), *The grammar of Romanian*, 24–64. Oxford: Oxford University Press.
- Zukoff, Samuel. 2015. Stress Restricts reduplication. *Proceedings of the Annual Meetings on Phonology* 2014. Washington, DC: Linguistic Society of America.

Dative clitics in Romanian ditransitives

Alina Tigău and Klaus von Heusinger

University of Bucharest / University of Cologne

This chapter presents a novel analysis for Romanian ditransitives. Based on empirical findings, we develop a derivational account building on the internal make-up of the two internal arguments. The account departs from the observation that clitic doubling (CD) of indirect objects (IO) and Differential Object Marking (DOM) of direct objects (DO) interact in an interesting and unexpected way: while unmarked DOs bind IOs irrespective of Clitic Doubling of IOs, counterparts where DOM-ed DOs bind CD-ed IOs are degraded; however, CD-ed+DOM-ed DOs fare much better. These facts seem to arise as a consequence of the interaction between DOM-ed DOs and CD-ed IOs which have similar internal make-up (they both carry a [Person] feature) and compete for the same probe, with the closer blocking agreement of the other. When DO cliticizes, these intervention effects no longer arise.

Keywords: ditransitives, differential object marking, clitic doubling

1. Introduction

One of the important aspects concerning ditransitive configurations concerns the relatedness between the ‘Prepositional Object Construction’ (POC) (1a) and the ‘Double Object Construction’ (DOC) (1b):

- (1) a. *Mary gave a book* Theme *to John* Goal. (POC)
b. *Mary gave John* Goal *a book* Theme. (DOC)

According to some studies, the two configurations are independent from one another, while others stress their structural connectedness. The former approach is known as the Alternative Projection Account, while the latter bears the name of Derivational Account.¹

1. Proponents of the Alternative Account include Oehrle (1976); Marantz (1993); Pesetsky (1995); Harley (2002); Bruening (2001, 2010). Advocates of the Derivational Account include Dryer (1987); Larson (1988, 1990); Baker (1988, 1996); Aoun & Li (1989); den Dikken (1995); Ormazabal & Romero (2010, 2012).

Proponents of the former account justify their claims as to the lack of relatedness between the two constructions by drawing on two arguments: the ‘lack of semantic uniformity’ of the alternating dative constructions (Oehrle 1976; Kayne 1975, among others) and the ‘asymmetric binding’ potential exhibited by the two internal arguments (Barss & Lasnik 1986; Aoun & Li 1989). The analyses advanced assume that the semantic differences holding between the two configurations are systematic, with POC expressing obligatory caused movement and the DOC describing obligatory caused possession. This difference has been accounted for in various ways, either in terms of a different event structure (Krifka 2004), as springing from the preposition relating the two internal arguments (Harley 2002) or as a difference between the light verbs *go* and *cause* present in the POC and the DOC, respectively (Cuervo 2003, among others).

For the proponents of the Derivational Account, the two configurations are related, with one of them representing the basic structure, and the other its syntactically derived counterpart. The thematic structure of V is argued to be the same in both frames and the semantic differences are explained in terms of affectedness deriving from the position occupied by IO. More recent work dwelling on idiomatic expressions, possession restrictions or inference patterns showed that differences in interpretation may actually remain semantically unexpressed, given that there is no strict correspondence between meaning and syntactic structure (Ormazabal & Romero manuscript, 2007; Rappaport Hovav & Levin 2008, among others). Rappaport Hovav & Levin (2008) show, for instance, that the ‘caused motion’ meaning associated with POC is not present with some of the verb classes exhibiting dative alternation and allowing POC (e.g., verbs of future having). Moreover, the ‘caused possession’ reading, linked to DOC, may surface with POC of certain verb classes expressing caused motion.

One important aspect extensively discussed by both approaches has to do with the *c*-commanding potential that the two internal arguments exhibit in the two configurations: in the POC configuration the DO *c*-commands the IO (2), but in the DOC structure the opposite *c*-command relation obtains (3).

POC: *Theme c-commands Goal*

- (2) a. *I showed Mary_i to herself_i (in the mirror).*
 b. **I showed herself_i to Mary_i (in the mirror).*

DOC: *Goal c-commands Theme*

- (3) a. *I showed John_i himself_i (in the mirror).*
 b. **I showed himself_i John_i (in the mirror).* (Barss & Lasnik 1986, 347 ex. 2)

Numerous studies on Romance ditransitives have adopted the Alternative Projection account, assuming structural and semantic differences between ditransitives

containing CD-ed IOs and their undoubled counterparts by grouping the former with DOC and analyzing the latter as POCs (Demonte 1995; Cuervo 2003).

Romanian was also argued to pattern with other Romance languages in this respect. Diaconescu & Rivero (2007) distinguish between ditransitives containing undoubled IOs (4a) (which they range with the English POC), and configurations where a dative clitic doubles IO (4b) (which they assimilate to the English DOC):

- (4) a. *Mihaela trimite Mariei o scrisoare.*
 Mihaela sends Mary.DAT a letter
 “Mihaela sends a letter to Mary.”
- b. *Mihaela îi trimite Mariei o scrisoare.*
 Mihaela CL.DAT.SG sends Mary.DAT a letter
 “Mihaela sends Mary a letter.” (Diaconescu & Rivero 2007: 210 ex. 1,2)

They further argue that the differences regarding *c*-command relations in the two configurations uncovered by Barss & Lasnik (1986) hold for Romanian ditransitives as well: in the configuration containing an undoubled IO (corresponding to POC) the DP_{Theme} is argued to *c*-command the DP_{Goal} (5a), while in the configuration featuring a clitic doubled IO, the opposite *c*-command relation is claimed to hold (5b). Note further that this latter structure is also posited to contain an Applicative Projection taking VP as its complement (Pylkkänen 2002):

- (5) a. POC: *Theme c-commands Goal*
 [VoiceP DP_{Agent} Voice_{[vP V [PP DP_{Theme} P DP_{Goal}]]]]}
- b. DOC: *Goal c-commands Theme* (clitic doubling)
 [VoiceP DP_{Agent} Voice_{[vP V [AppIP DP_{Goal} [cl_{Appl}] [VP V DP_{Theme}]]]]]}
- (Diaconescu & Rivero 2007: 219–220)

The Appl head in DOC is occupied by the dative clitic, following a close parallelism with Cuervo’s (2003) proposal for Spanish. The two different configurations are thus triggered by the presence of the dative pronominal clitic doubling IO or the lack thereof: in (5a), the undoubled IO merges low within PP, while DO occupies Spec,PP, *c*-commanding IO. In (5b), the clitic doubled IO merges in Spec,Appl, while DO occupies the complement position. Appl⁰ spells out as the dative clitic. As such, the binding asymmetries between the two arguments pattern with the ones in English.

The account put forth by Diaconescu & Rivero makes a number of predictions in the sense that some configurations are discarded as ungrammatical, while others are predicted to be grammatical: DOs may not bind CD-ed IOs since the latter DP merges in a *c*-commanding position (6b); DO may only bind an undoubled IO, given its low position within the PP (6a):

- (6) a. *Poliția a dat tatălui său, copilul pierdut.*
 Police.the has given father.DAT his child.the lost
 “The police gave the lost child to his father.”
- b. ^{??}*Poliția i-a dat tatălui său, copilul pierdut.*
 Police.the CL.DAT.SG-has given father.DAT his child.the lost.
lit. “The police gave his father the lost child.”
- (Diaconescu & Rivero 2007: 223–224 ex. 28b,30b)

On the other hand, an undoubled IO is unable to bind into DO, since it merges in a lower position (7b); the clitic-doubled IO is, on the other hand, able to bind into the IO (7a).

- (7) a. *I-am dat muncitorului, cecul său.*
 CL.DAT.SG-have.I given worker.DAT cheque.the his
 “I have given the worker his cheque.”
- b. ^{*?}*Am dat muncitorului, cecul său.*
 have.I given worker.DAT cheque.the his
lit. “I gave the worker his cheque.”
- (Diaconescu & Rivero 2007: 223–224 ex. 28a, 30a)

More recently, Cornilescu et al. (2017a) showed that language facts do not match these theoretical expectations and emphasized the availability of symmetric c-command within ditransitives.

This contribution builds on this idea and extends the analysis in order to capture the complex interaction between the two internal arguments within ditransitive configurations. Romanian is quite complex in this respect in that it allows marking of its DOs by means of the differential marker *pe*, which is sensitive to the animacy and definiteness scales (Aissen 2003; Tigău 2011), and by means of an accusative pronominal clitic.

Our account focuses particularly on marked direct objects, i.e., single differentially object-marked DOs (DOM-ed DOs) and clitic-doubled and differentially object-marked DOs (CD-ed+DOM-ed DOs) when these co-occur with clitic-doubled IOs. Special attention is granted to those configurations featuring DOM-ed DOs and CD-ed IOs, which have turned out to be problematic with respect to their acceptability for native speakers of Romanian.

The data addressed in this chapter and for which we are proposing an analysis have been gathered as part of a series of experiments checking binding relations between the two internal arguments within Romanian ditransitives (§ 2).

Note also that Diaconescu & Rivero (2007) do not make any predictions with respect to binding relations when marked direct objects are involved. For a detailed discussion along these lines see Tigău (2020); Tigău & von Heusinger (manuscript).

The chapter is structured as follows: § 2 contains a discussion of some binding problems with ditransitives uncovered experimentally and the predictions we start

from; § 3 discusses the feature specification of DO and IO; § 4 provides an account for the two problems experimentally uncovered; § 5 contains the conclusions. The three experiments are presented in the appendix.

2. One problematic configuration

In the following, we present original data on the binding properties of the internal arguments of ditransitive configurations from a broad empirical survey. In three grammaticality judgment tasks we manipulated word order ('DO before IO' vs 'IO before DO'), binding direction ('DO binds into IO' vs 'IO binds into DO') and clitic doubling of the IO. Thus, each experiment consisted of a $2 \times 2 \times 2$ design. Between the three experiments we varied the layout of the direct object: in Expt 1 we used inanimate DOs and therefore unmarked DOs, while in Expt 2 and Expt 3 human DOs were employed. Expt 2 and Expt 3 differed from one another in that, while Expt 2 featured single DOM-ed DOs, Expt 3 drew on CD-ed+DOM-ed DOs. 32 sentences were designed for each experiment and varied, changing word order, binding direction and presence/absence of the dative clitic so that we had 256 items for each experiment, which were distributed into 8 lists using the Latin square method. 32 fillers were added. Each list in each experiment was assessed by at least 20 native speakers: more than 160 people participated in each experiment.

In this chapter, we discuss only one part of the full results, namely the condition 'DO before IO and DO binds IO' and vary the clitic doubling of the IO and the layout of the DO, as in Table 1 (for a full presentation of the data, the results and an analysis, see Tigău (2020) and Tigău & von Heusinger (manuscript)).²

2. Table 1 is an abstraction from the absolute figures in Table 2 with the mean values of the grammaticality judgments from 1 (very bad) to 7 (very good).

Table 2. Mean values of acceptability of binding configuration of DO > IO and DO binds into IO with different forms of DO and undoubled vs CD-ed IO (see Tigău (2020) for full information)

	IO	CD-ed IO
<i>unmarked DO</i>	4,57	3,64
<i>DOM-ed DO</i>	4,43	2,64
<i>CD-ed+DOM-ed DO</i>	4,51	3,52

The difference for DOM-ed DOs and undoubled IOs (4,43) vs CD-ed IOs (2,64) is significant: Statistical analysis was conducted in R version 1.0.136 using the lme4 package (Bates et al. 2014) to perform linear mixed-effect models (LMEM) with the score as outcome variable. As fixed effects, we entered word order, Binding and Clitic Marking into the model. As random effects, we had intercepts for subjects and items. The word order 'DO before IO

Table 1. Acceptability (++ very good, + acceptable, – bad) of binding configuration DO > IO and DO binds into IO with different types of DO (unmarked, DOM-ed, CD-ed and DOM-ed) and IO (undoubled, CD-ed) extracted from 3 questionnaires with 120 informants each

	IO	CD-ed IO
<i>unmarked DO</i>	++ cf. (8a)	+ cf. (8b)
<i>DOM-ed DO</i>	++ cf. (9a)	– cf. (9b)
<i>CD-ed+DOM-ed DO</i>	++ cf. (10a)	+ cf. (10b)

One of the facts uncovered experimentally concerns the low acceptability of ditransitives containing DOM-ed DOs and CD-ed IOs in the DO before IO word order. This seemed unusual given that, when compared with ditransitives containing unmarked DOs or CD-ed+DOM-ed DOs, these instances fared significantly worse.

Example (8) shows that unmarked DOs may bind both undoubled IOs and CD-ed IOs. (9) shows that while DOM-ed DOs may bind the possessor within undoubled IOs, the same configuration is sharply degraded when IOs are CD-ed. (9b) is saved if the DOM-ed DO is CD-ed, (10a).

(8) unmarked DO > undoubled IO

- a. *Editorii au trimis fiecare carte_i autorului ei_i pentru*
 editors.the have sent every book author.DAT its for
corecturile finale.
 corrections final

“The editors send each book to its author for the final corrections.”

unmarked DO > CD-ed IO

- b. *Editorii i-au trimis fiecare carte_i autorului ei_i pentru*
 editors.the CL.DAT.SG-have sent every book author.DAT its for
corecturile finale.
 corrections final

“The editors send each book to its author for the final corrections.”

(9) DOM-ed DO > undoubled IO

- a. *Comisia a repartizat pe fiecare medic rezident_i unor foști*
 board.the has assigned DOM every doctor resident some.DAT former
profesori de-ai lui_i.
 professors of his

“The board assigned every medical resident to some former professor of his.”

condition’, Binding ‘DO binds into IO’ condition and the Clitic Marking ‘no clitic’ condition were mapped onto the intercept. To identify the best model fit we performed likelihood ratio tests. This revealed that the full model with a three-way interaction affected the acceptance rate ($\chi^2(4) = 36.21, p < .001$).

DOM-ed DO > CD-ed IO

- b. **Comisia le-a repartizat pe fiecare medic*
 board.the CL.DAT.PL-has assigned DOM every doctor
rezident_i unor foști profesori de-ai lui_i.
 resident some.DAT former professors of his
 “The board assigned every medical resident to a former professor of his.”

(10) CD-ed+DOM-ed DO > undoubled IO

- a. *Comisia l-a repartizat pe fiecare*
 board.the CL.ACC.SG.M-has assigned DOM every
medic rezident_i unor foști profesori de-ai lui_i.
 doctor resident some.DAT former professors of his
 “The board assigned every medical resident to some former professor of his.”

CD-ed+DOM-ed DO > CD-ed IO

- b. *Comisia li l-a repartizat pe fiecare medic*
 board.the CL.DAT.PL CL.ACC.SG.M-has assigned DOM every doctor
rezident_i unor foști profesori de-ai lui_i.
 resident some.DAT former professors of his
 “The board assigned every medical resident to some former professor of his.”

The aim of this paper is to propose an account which would accommodate the differences between unmarked DOs and their DOM-ed counterparts. More specifically, two questions will be addressed:

1. Why is the co-occurrence of DOM-ed DOs and CD-ed IOs assessed as unacceptable, while configurations containing unmarked DOs and CD-ed IOs fare quite well?
2. Why does CD of the DO improve the acceptability of configurations with DOM-ed DOs and CD IOs?

One way to approach these facts would be to start by considering the following: given that configurations with unmarked DOs and CD-ed+DOM-ed DOs fare similarly with respect to acceptability scores and seem to be felicitous, we should not hold the binding dependency itself to be responsible for the low acceptability of counterparts with DOM-ed DOs. What seems to be the problem is the co-occurrence of DOM and the dative clitic doubling the IO. This might indicate that the lower acceptability of these instances has to do with the internal structure of the object DPs involved. The next sections propose an account along these lines.

3. The featural make-up of IOs and DOM-ed DOs

Romanian aligns with other DOM languages and may differentially mark its DO by means of the marker *pe*, a derivative of the locative preposition *pe* (“on”). Cornilescu (2000, among others) views DOM primarily as a marker of ‘semantic gender’ used with person denoting DPs and disallowed with non-person-denoting ones. Other studies connect DOM with animacy with definiteness (Aissen 2003, among others), others with specificity (Farkas 1987; Dobrovie-Sorin 1994; von Heusinger & Tigău 2019, among others). In the next section we try to capture the contribution of DOM by building on the notion of [Person] (Richards 2008).

3.1 A [iPerson] for DOM-ed DOs

The idea that DOM-ed DOs bear a [Person] feature is not new (e.g., for Spanish Mondoñedo 2007, among others) and is in line with the sensitivity of these DPs to the animacy and definiteness hierarchies (Silverstein 1986; Aissen 2003). Romanian DOM-ed DOs were argued to pattern with their Spanish counterparts. Cornilescu (2000) argues that *pe* represents a mark of ‘personal’ gender and ‘identification’ and justifies her claim by building on the behaviour of bare quantifiers *nimeni* (“nobody”), *cineva* (“somebody”), which always get *pe* when used as DOs. As such, there is a clear-cut distinction between bare quantifiers referring to persons and those referring to non-persons:

- (11) *N-am văzut *(pe) nimeni/(*)pe nimic.*
 not-have.I seen DOM nobody/ DOM nothing
 “I haven’t seen anybody.”

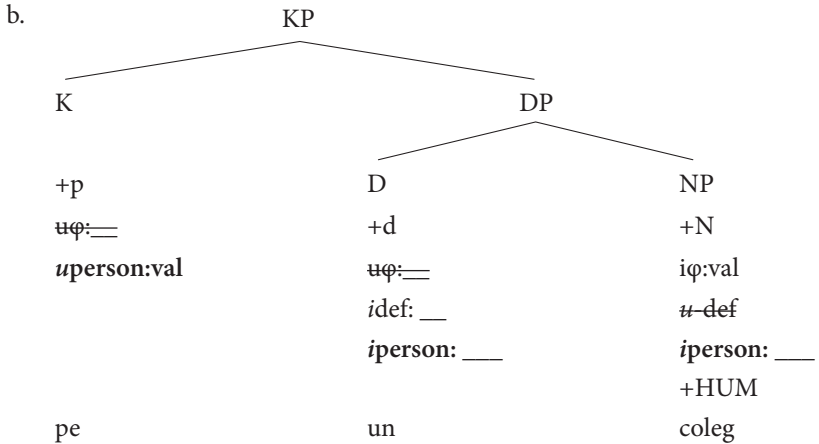
Cornilescu (2000) proposes that *pe* is a means of expressing ‘semantic gender’, i.e., a notion distinguishing between non-neuter gender (personal) and neuter (non-personal) gender. [–Neuter] DPs are semantically marked as [+Person] and require *pe*, while [+Neuter] DOs are semantically marked as [–Person] and will not require *pe*.

Building on this idea and following López (2012), as well as Cornilescu & Tigău (manuscript), we assume that the internal makeup of marked DOs presupposes the existence of a KP layer, where K is triggered by an unvalued syntactic [iPerson] feature present in the NP (i.e., [iPerson:___] (12b)). The [iPerson:___] is then copied in D.³ The NP itself is a [+Human] denoting nominal and as such may incorporate the [iPerson] feature. The presence of the syntactically unvalued [iPerson] feature

3. We adopt Pesetsky & Torrego (2007) as a general framework for feature checking.

triggers the merger of K (*pe*) which carries a valued [*uPerson*]. The valued [*uPerson*] feature on K checks the unvalued feature [*iPerson*:__] on D. As a consequence, the entire KP ends up bearing a valued [*iPerson*: val]:

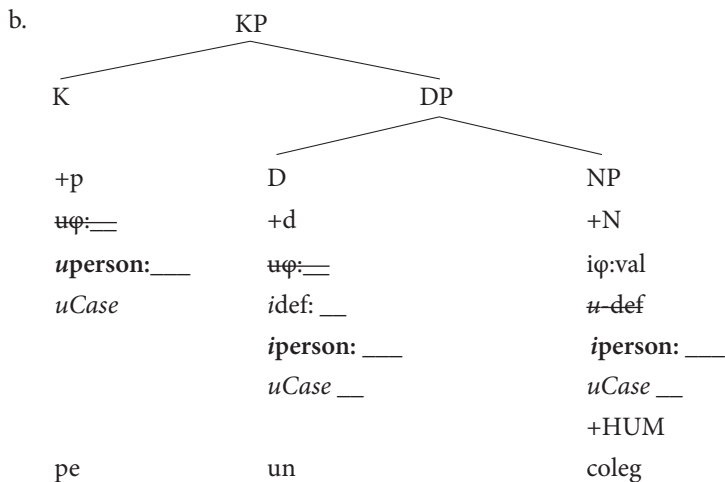
- (12) a. *Ajut pe un coleg.*
 help.I DOM a colleague
 ‘‘I help a colleague.’’



The case of CD+DOM-ed DOs is somewhat different given that the marker *pe* has been shown to function differently than its counterpart in single DOM-ed DOs.⁴ We capture this difference by positing that in the case of CD+DOM-ed DOs, *pe* carries an ‘unvalued’ [*uPerson*:__] feature. Thus, the ‘Person’ feature on K is in need of valuation, like the one on the nominal it precedes (which, just as above, bears [*iPerson*:__]). After agreement between K and DP applies, KP ends up bearing an unvalued [*iPerson*: __] feature and has to find a way whereby to value it (13b):

- (13) a. *Îl ajut pe un coleg.*
 CL.ACC.SG.M help.I DOM a colleague
 ‘‘I help a colleague.’’

4. Avram (2014) relies on the results of an experiment involving two acceptability sentence questionnaires from 23 native speakers of Romanian (age 20–57). The experiment proved that the participants fall into two categories: those that always clitic double the *pe* marked DP and those who allow for ‘single’ *pe* marking besides having the option of CD+*pe*. Based on these results, Avram (2014) speaks of two instances of *pe*: the ‘single’ *pe*, as a semantic gender marker (as described by Cornilescu 2000) and the *pe* in CD structures (which she takes to be an accusative case marker).



Following Ciucivara (2009), we posit a PersonP at the νP periphery: CD-ed+DOM-ed DOs undergo scrambling and have an [*i*Person __] feature valued under agreement with Pers⁰. Single DOM-ed DOs, on the other hand, will not need to move into this position, given that their ‘person’ feature is a valued, interpretable one ([*i*Person: val], see (12b)).

3.2 A [Person] feature for Goal DPs

As shown by Tigău (2014, among others), the essential property of DPs which may realize the dative theta roles is sensitivity to the animacy hierarchy. These roles seem to denote human individuals, i.e., DPs marked for [+Person]:⁵

(14) Possessor – Goal

Profesorul le-a înapoiat tezele elevilor/ la elevi.
 professor.the CL.DAT.PL-has returned theses.the pupils.DAT to pupils
 “The professor returned the theses to the pupils.”

(15) Beneficiary

Mama i-a cusut rochia fetei/ la fată.
 mother CL.DAT.SG-has sewn dress.the girl.DAT to girl
 “Mother has sewn the dress for the girl.”

5. Romanian datives exhibit ‘inflectional’ or ‘prepositional’ case marking. Prepositional marking presupposes the use of the directional preposition *la* (“at/to”) and is used with DPs headed by invariable determiners. e.g., *niște* (“some”), cardinals, etc. As argued in Diaconescu & Rivero (2007), this variant seems to be more frequently employed in the north-western part of the country, while the inflectional form is preferred in educated Romanian.

- (16) Maleficiary
Copiii le-au furat vecinilor/ la vecini
 children CL.DAT.PL-have stolen neighbours.DAT to neighbours
cireșele din grădină.
 cherries from garden
 “The children stole the neighbour’s cherries from the garden.”

Note that IO does not usually denote inanimate referents, at least when used in the inflectional dative:⁶

- (17) *Am dat apă florilor/ la flori.*
 have.I given water flowers.DAT to flowers
 “I watered the flowers.”

6. One reviewer correctly points out that Romanian marginally allows certain inanimate datives providing the following examples:

- (i) a. *Am pus zahăr cafelei.*
 have.I put sugar coffee.DAT
 “I have put sugar into the coffee.”
 b. *A dat un șut scaunului de a zburat pe fereastră.*
 has given a kick chair.DAT that has flown on window
 “He kicked the chair out of the window.”

We consider these examples marginal indeed, with a preference for the prepositional dative:

- (ii) a. *Am pus zahăr la cafea.*
 Have.I put sugar to coffee.
 “I have put sugar into the coffee.”
 b. *A dat un șut la scaun de a zburat pe fereastră.*
 has given a kick to chair that has flown on window
 “He kicked the chair out of the window.”

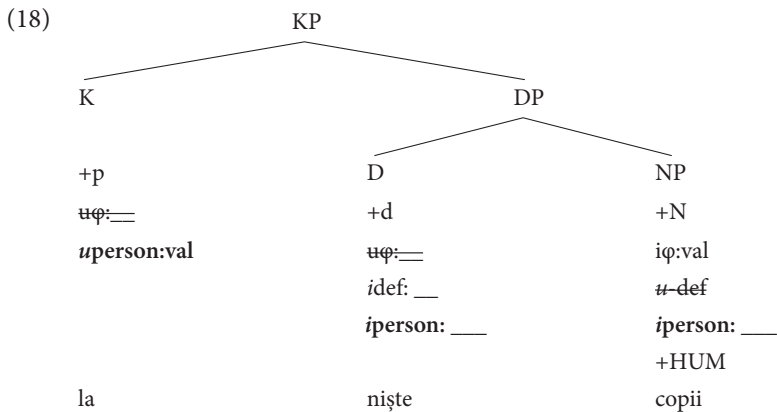
Acceptance of examples in (i) might show a process of unification of uses of the inflectional dative and the prepositional one, considering that the latter allows marking of [–human] IOs.

Note, on the other hand, that abstract nouns may function (metaphorically) as inflectional datives. In this case, the prepositional dative is disallowed:

- (iii) *A supus proiectul atenției bordului /^{*}la atenția bordului.*
 has.he submitted project.the attention.the.DAT board.the.GEN to attention.the board.the.GEN
lit. “He submitted the project to the board’s attention.”

A further interesting phenomenon concerns clitic doubling of the IO in these examples: the reviewer accepts doubling in (ib) but rejects it in (ia), an intuition which we also agree with. Note, however, that when using the prepositional dative (in ii), doubling becomes possible for both variants. In (iii), on the other hand, doubling is out. We leave this matter for further research.

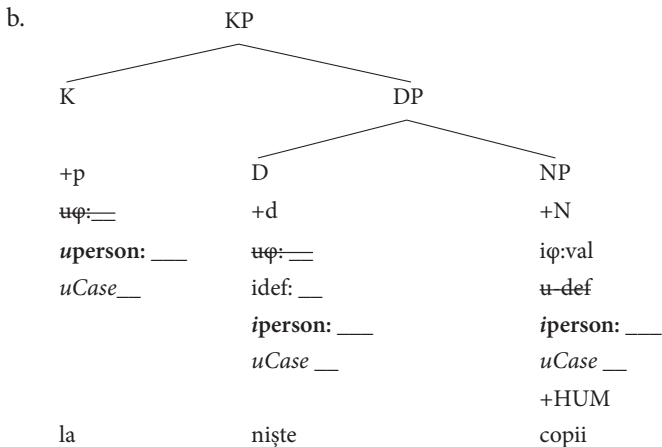
We will capture this sensitivity by positing that dative DPs carry a [Person] feature, like DOM-ed DOs. These nouns grammaticalize their animacy feature as a [Person] feature (Richards 2008). Just as above, we posit that [+Human(like)] NPs incorporate an unvalued syntactic [*i*Person:___] feature, which is copied in D and checked by merging K. K itself carries a valued [*u*Person: val] feature. (18) shows this at work: the existence of an unvalued feature in N and then in D triggers the insertion of K. After agreement between K and DP, the entire KP has the feature specification [*i*Person: val].



Inflectional datives will follow the same pattern of analysis and evince the same feature specifications as KP, possessing a silent K head.

Clitic-doubled IOs will closely pattern with CD-ed+DOM-ed DOs (13): like the differential marker *pe* for DOs, the K on doubled dative DPs will carry an unvalued [Person] feature, i.e., [*u*Person:___]. The results of feature agreement between K and DP this time is an unvalued [*i*Person: ___] (19b). This feature will thus have to be further checked in the course of the derivation. As such, there is complete parallelism between CD-ed+DOM-ed DOs and CD-ed IOs with respect to feature checking:

- (19) a. *Le-am dat cartea la niște colegi.*
 CL.DAT.PL-have.I given book.the to some colleagues.
 “I gave the book to some colleagues.”

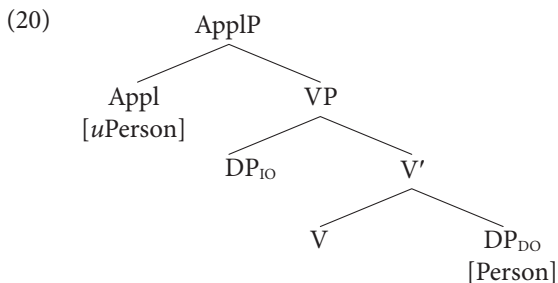


Doubled inflectional datives pattern on a par with their prepositional counterparts, exhibiting a silent K.

Note also that undoubled IOs pattern with DOM-ed DOs with respect to their feature specification, both expressions being KPs and bearing an [*i*Person: val] feature specification. Drawing the parallelism even further, we have assimilated CD-ed+DOM-ed DOs with CD-ed IOs with respect to their status as KPs and their feature specification as [*i*Person: ____]. In this latter case, the dative K only contributes an unvalued [*u*Person: ____] feature, which does not suffice to value [*i*Person: ____] carried by the nominal expression. The feature of KP will thus have to be valued at a later point during the derivation (most probably against Appl⁰ of the Applicative Projection proposed for ditransitives, see Pylkkänen 2002, 2008, among others, in this sense).

Drawing on Marantz (1993) and Pylkkänen (2002), we posit an Applicative projection for ditransitives (20). In line with Georgala et al. (2008) and Georgala (2012), we envisage ApplP as a case assigner also introducing a [Person] feature thereby capturing the sensitivity of datives to the animacy hierarchy. The [Person] feature accounts for the variety of theta roles compatible with dative DPs within ditransitives given that all these roles presuppose the presence of a [+Human] feature (see above).

Thus, ApplP takes VP as its complement (20) and introduces an unvalued uninterpretable [Person] feature, which may be checked by way of agreement with the dative DP also carrying a [Person] feature:



4. A syntactic account of the experimental data

As already pointed out, the primary goal of the current chapter is to account for the experimentally noticed differences between unmarked DOs and their DOM-ed counterparts when co-occurring with CD-ed IOs: while co-occurrence of DOM-ed DOs with CD-ed IOs is discarded as unacceptable, ditransitives with unmarked DOs and CD-ed IOs fare quite well with respect to acceptability judgments. A second aim of this article is to provide an explanation for the acceptability judgments regarding structures where CD-ed+ DOM-ed DOs co-occur with CD-ed IOs, assessed as acceptable. In other words, we need to explain why CD of DOM-ed DO functions as a repair strategy, given that the co-occurrence of undoubled DOM-ed DOs and CD IOs is out.

In what follows, we address these two questions by building on the initial intuitions presented in Cornilescu et al. (2017b). More specifically, we adopt a derivational account according to which dative DPs are merged within the VP as part of the verb's argument structure. In line with Larson (2010)'s view, IO is 'actually part of the verb's θ -grid'. It is introduced by the lexical verb itself and composes inside VP in a syntax similar to that in Larson (1988). Under this view, Appl⁰ is required to have the lower lexical VP as complement.

Also, as discussed in the last section, marked DOs and IOs will bear a [Person] feature, which is further specified as 'interpretable/uninterpretable' and as 'valued/unvalued' function of various factors as described above: undoubled IOs and DOM-ed DOs carry a [*i*Person: val] feature specification, while CD-ed+DOM-ed DOs and CD-ed IOs are specified as [*i*Person: ____].

We further propose that there is a certain priority regarding feature verification between the two objects.⁷ More specifically, the DO will have priority over the IO.

7. We found this priority requirement (springing from the need of feature valuation) crucial when considering the derivation of all the available tested configurations. See Tigău (2020) for further clarifications.

Priority may, however, change function of the feature specification of the two objects. The following cases arise:

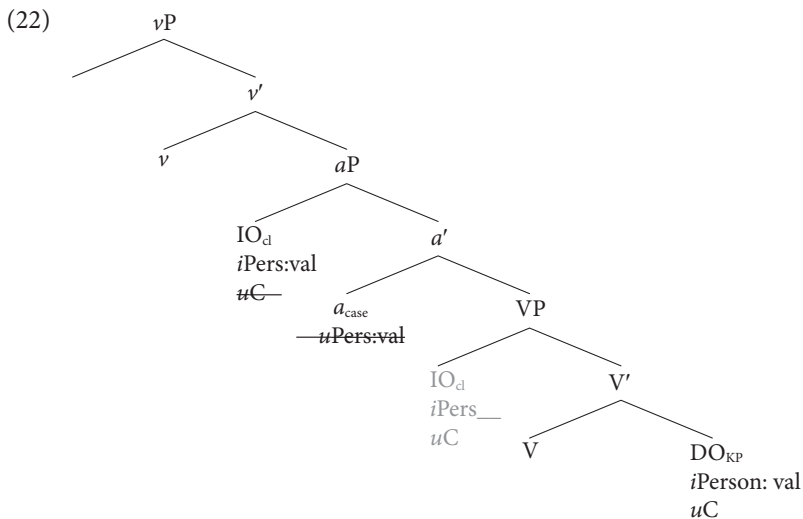
1. Unmarked DOs only bear [uC] and have no specification with respect to [Person] in syntax; IO will always have both [uC] and [Person] (irrespective of whether this latter feature is [i Person: val] as with undoubled IOs or [i Person: ___] as with their doubled counterparts). Given that DO has no [Person] feature to verify, it will simply undergo scrambling first.
2. DOM-ed DOs bear [uC] and [i Person: val]. In this case, both DO and IO are sensitive to [Person] so a prioritization as to which of them values their [Person] feature first needs to occur. Two situations may arise:
 - a. IO has the same feature specification, i.e., [uC] and [i Person: val] (as it is undoubled): DO will be given priority for movement.
 - b. IO is doubled and as such has more features to verify, i.e., [uC] and [i Person: ___]: in this particular case, the IO will gain priority over the DO, which only needs to verify one feature.
3. CD-ed+DOM-ed DOs bear [uC] and [i Person: ___] and will always have priority over the IO:
 - a. If the IO is undoubled its feature specification is [uC] and [i Person: val] > DO has priority because of its DO status and also because it will have more features to verify.
 - b. If the IO is doubled, then it will have the same feature specification as the DO, i.e., [uC] and [i Person: ___] > DO has priority according to the initial criterion

4.1 Unmarked DOs and CD-ed IOs

Configurations where an unmarked DO co-occurs with a CD-ed IO fare well with respect to acceptability judgments expressed by our respondents. Consider first the example und (21):

- (21) *Editorii i-au trimis fiecare carte_i autorului ei_i pentru*
 editors.the CL.DAT.SG.-have sent every book author.DAT its for
corecturile finale.
 corrections final
 ‘‘The editors sent every book to its author for the final corrections.’’

In this particular case, only IO evinces sensitivity to [Person] and bears [i Person: ___]. IO also carries [u Case]. DO has no [Person] specification and only needs to check case. According to the priority criteria adopted above, DO will thus be the first to enter the derivation and move into the specifier of α P where it values its case feature against v . IO will verify both case and [Person] against the α head:



The derivation above explains both directions of binding in the ‘DO before IO’ word order. The ‘IO before DO’ obtains by scrambling the IO to a specifier of the νP : in this way IO reaches a c-commanding position with respect to the DO and both directions of binding find an explanation.

4.2 DOM-ed DOs and CD-ed IOs

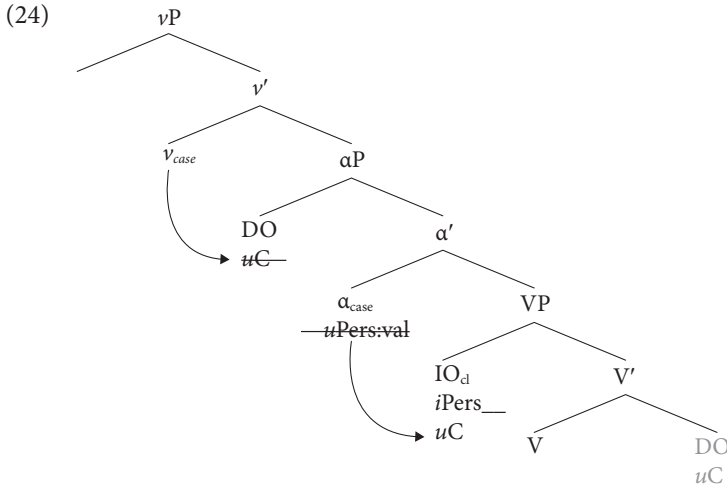
According to the experimental findings, sentences such as (23) were granted very low acceptability scores by the respondents:

- (23) **Delegații i-au lăudat pe fiecare secretară_i șefului ei_r.*
 delegates.the CL.DAT.SG-have praised DOM every secretary boss.DAT her
lit. “The delegates have praised every secretary to her boss.”

Consider this derivation at work: we start from VP where the DO is merged in the complement position and has the feature specification: [*iPers: val*], [*uCase*]. DO thus only needs to verify its [*uCase*] feature, given that its [*Person*] feature is both interpretable and valued. IO, on the other hand, carries an unvalued [*iPerson*] feature along with [*uCase*] and will have to find a way to value both these features.

Note that both objects are specified for Person, but that IO has more features to verify and will gain priority over DO. IO enters Agreement with α^0 (specified as [*uPerson:val*]) and checks both case and [*iPers:*]. The [*uPerson:val*] feature of α is EPP and IO moves to Spec, αP . As such, it acts as an intervener for DO, which may no longer move to a Spec, αP in order to get its case valued by ν (24).

Thus, movement of DO out of VP is not possible, hence the derivation crashes. This explains the low results in the ‘DO before IO’ word order: DO may not leave the VP. One way to save the situation is by scrambling IO out of Spec, α P, into a specifier of v . As a consequence, IO will no longer act as an intervener for DO, which may scramble to a specifier of α and get its case feature valued by v . This explains why the order IO before DO was found to be significantly better than its opposite.



4.3 CD-ed + DOM-ed DOs and CD-ed IOs

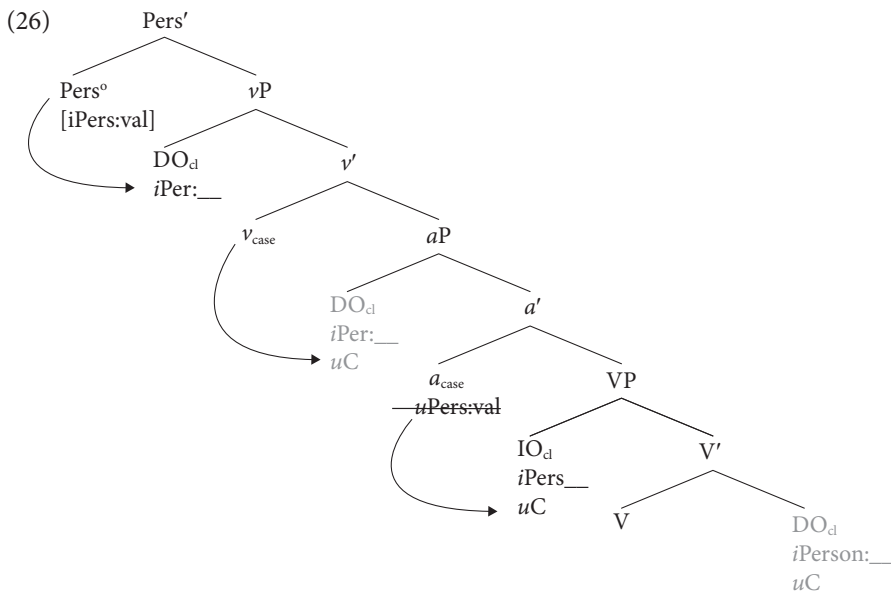
As already observed, these configurations fare much better with respect to acceptability judgments as opposed to their counterparts containing undoubled DOM-ed DOs which were assessed as thoroughly degraded:

- (25) *Delegații i-au lăudat-o pe fiecare secretară,*
 delegates.the CL.DAT.SG-have praised-CL.ACC.SG.F DOM every secretary
șefului ei,
 boss.DAT her
 ‘The delegates have praised every secretary to her boss.’

In this particular situation, DO and IO have the same feature specification: [i Person:___] and [u Case]⁸ and therefore DO has priority over IO. As a consequence, it will move to Spec, α P and value its case feature against v . Given that DO also

8. See above on how we get to the feature specifications of DO and IO with respect to Person.

needs to value [*i*Person:___], this KP moves further, to a specifier of *v* and enters agreement with the Person⁰. IO will be probed by the α^0 and will have [*i*Person:___] and case valued against this head. The [*u*Person: val] of a will also be checked as a consequence:



This derivation shows that both binding directions are possible, given that DO may occupy a position wherefrom it may c-command IO and the other way round. The IO before DO word order may be easily obtained by scrambling IO to a specifier of *v*.

5. Conclusions

This paper has provided an answer to two questions arrived at experimentally: the first question revolved around the infelicity with respect to acceptability of configurations wherein DOM-ed DOs co-occur with CD-ed IOs, while a further question concerned the role of the accusative clitic doubling DOM-ed DOs when present in the same configurations – as it seems, CD of DOM-ed DO functions as a repair strategy in this case. With respect to the former question, it was argued that the interaction between DOM-ed DOs and CD-ed IOs boils down to a locality issue: VAppl, which may match both nominals in its c-commanding domain in what the valuation of its [*u*Person] feature is concerned, may only do so with the higher object, in our case IO.

When DO cliticizes, however, it will have priority over IO and will be able to leave the VP and check its case feature against *v*. It will move further into a position where it will be able to also check [Person] against Person⁰. IO will enter agreement with Appl⁰ and thereby check case and Person.

Acknowledgements

We would like to thank the audience of Going Romance in Bucharest, December 2017, for their critical comments, two anonymous reviewers for helpful comments, and Elyesa Seidel for providing us with the statistics. We also thank Adina Dragomirescu and Alexandru Nicolae for their support and for editing this volume.

Funding

The research for this paper has been funded by the Alexander von Humboldt Foundation for the first author and by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) – Projektnummer 281511265 – SFB “Prominence in Language” in the project B04 “Interaction of nominal and verbal features for Differential Object Marking” at the University of Cologne for the second author.

References

- Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language & Linguistic Theory* 21. 435–483. <https://doi.org/10.1023/A:1024109008573>
- Aoun, Joseph & Audrey Yen-Hui Li. 1989. Scope and constituency. *Linguistic Inquiry* 20. 141–172.
- Avram, Larisa. 2014. Differential object marking in Romanian: The view from acquisition. Paper presented at the 16th Annual Conference of the English Department (ACED 16), Bucharest, June 6–8.
- Baker, Mark. 1988. Theta theory and the syntax of applicatives in Chichewa. *Natural Language & Linguistic Theory* 6. 353–389. <https://doi.org/10.1007/BF00133903>
- Baker, Mark. 1996. On the structural position of themes and goals. In Johan Rooryck & Laurie Zaring (eds.), *Phrase structure and the lexicon*, 7–34. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-015-8617-7_2
- Barss, Andrew & Howard Lasnik. 1986. A note on anaphora and double objects. *Linguistics Inquiry* 17. 347–335.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steve Walker. 2014. lme4: linear mixed-effects models using Eigen and S4. *R package version*, 1(7). 1–23.
- Bruening, Benjamin. 2001. QR obeys superiority: Frozen scope and ACD. *Linguistic Inquiry* 32(2). 233–273. <https://doi.org/10.1162/00243890152001762>

- Bruening, Benjamin. 2010. Double object constructions disguised as prepositional datives. *Linguistic Inquiry* 41(2). 287–305. <https://doi.org/10.1162/ling.2010.41.2.287>
- Ciucivara, Oana. 2009. A syntactic analysis of pronominal clitic clusters in Romance. Doctoral dissertation, New York University.
- Cornilescu, Alexandra. 2000. Notes on the interpretation of the prepositional accusative in Romanian. *Bucharest Working Papers in Linguistics* 2(1). 91–106.
- Cornilescu, Alexandra & Alina Tigău. Manuscript. *Landscaping datives*. University of Bucharest.
- Cornilescu, Alexandra, Anca Dinu & Alina Tigău. 2017a. Experimental data on Romanian double object constructions. *Revue roumaine de linguistique* 62(2). 157–177.
- Cornilescu, Alexandra, Anca Dinu & Alina Tigău. 2017b. Romanian dative configurations: Ditransitive verbs. A tentative analysis. *Revue roumaine de linguistique* 62(2). 179–206.
- Cuervo, Maria Cristina. 2003. Datives at large. Doctoral dissertation, MIT.
- Demonte, Violeta. 1995. Dative alternation in Spanish. *Probus* 7(1). 5–30. <https://doi.org/10.1515/prbs.1995.7.1.5>
- Diaconescu, Constanța Rodica & Maria Luisa Rivero. 2007. An applicative analysis of double object constructions in Romanian. *Probus* 19(2). 209–233. <https://doi.org/10.1515/PROBUS.2007.007>
- den Dikken, Marcel. 1995. *Particles: On the syntax of verb-particle, triadic, and causative constructions*. Oxford: Oxford University Press.
- Dobrovie-Sorin, Carmen. 1994. *The syntax of Romanian*. Berlin: de Gruyter Mouton. <https://doi.org/10.1515/9783110886597>
- Dryer, Matthew. 1987. On primary objects, secondary objects & antidative. *Language* 62. 808–845. <https://doi.org/10.2307/415173>
- Farkas, Donka. 1987. Direct and indirect object reduplication in Romanian. *Chicago Linguistics Society* 14. 88–97.
- Georgala, Effi. 2012. Applicatives in their structural and thematic function: A minimalist account of multitransitivity. Doctoral dissertation, Cornell University.
- Georgala, Effi, Waltraud Paul & John Whitman. 2008. Expletive and thematic applicatives. In Charles B. Chang & Hannah J. Haynie (eds.), *Proceedings of the West Coast Conference on Formal Linguistics*, 181–189. Somerville, MA: Cascadilla Proceedings Project.
- Harley, Heidi. 2002. Possession and the double object construction. In Pierre Pica (ed.), *Linguistic variation yearbook 2002*, 29–68. Amsterdam: John Benjamins. <https://doi.org/10.1075/livy.2.04har>
- von Heusinger, Klaus & Alina Tigău. 2019. Clitic doubling and differential object marking in non-specific contexts in Romanian. *Revue roumaine de linguistique* 66(4). 409–430.
- Kayne, Richard. 1975. *French syntax*. Cambridge, MA: MIT Press.
- Krifka, Manfred. 2004. Semantic and pragmatic conditions for the dative alternation. *Korean Journal of English Language and Linguistics* 4. 1–32.
- Larson, Richard K. 1988. On the double object construction. *Linguistic Inquiry* 19. 335–391.
- Larson, Richard K. 1990. Double objects revisited: Reply to Jackendoff. *Linguistic Inquiry* 21. 589–632.
- Larson, Richard K. 2010. On Pylkkänen's semantics for low applicatives. *Linguistic Inquiry* 41. 701–704. https://doi.org/10.1162/LING_a_00020
- López, Luis. 2012. *Indefinite objects. Scrambling, choice functions, and differential marking*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9165.001.0001>

- Marantz, Alec. 1993. Implications of asymmetries in double object constructions. In Sam A. Mchombo (ed.), *Theoretical aspects of Bantu grammar*, 113–150. Stanford, CA: CSLI Publications.
- Mondoñedo, Miguel Rodriguez. 2007. The syntax of objects: Agree and differential object marking. Doctoral dissertation, University of Connecticut.
- Oehrle, Richard. 1976. The grammatical status of the English dative alternation. Doctoral dissertation, MIT.
- Ormazabal, Javier & Juan Romero. Manuscript. *Agreement restrictions*. University of the Basque Country & University of Alcalá de Henares. 2002.
- Ormazabal, Javier & Juan Romero. 2007. The object agreement constraint. *Natural Language & Linguistic Theory* 25(2). 315–347. <https://doi.org/10.1007/s11049-006-9010-9>
- Ormazabal, Javier & Juan Romero. 2010. The derivation of dative alternations. In Maia Duguine, Susana Huidobro & Nerea Madariaga (eds.), *Argument structure and syntactic relations: A cross-linguistic perspective*, 203–232. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.158.13orm>
- Ormazabal, Javier & Juan Romero. 2012. PPs without disguises: Reply to Bruening. *Linguistic Inquiry* 43(2). 455–474. https://doi.org/10.1162/LING_a_00097
- Pesetsky, David. 1995. *Zero syntax: Experiencers and cascades*. Cambridge, MA: MIT Press.
- Pesetsky, David & Esther Torrego. 2007. The syntax of valuation and the interpretability of features. In Simin Karimi, Vida Samiian & Wendy K. Wilkins (eds.), *Phrasal and clausal architecture: Syntactic derivation and interpretation*, 262–294. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.101.14pes>
- Pylkkänen, Liina. 2002. Introducing arguments. Doctoral dissertation, MIT.
- Pylkkänen, Liina. 2008. *Introducing arguments*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9780262162548.001.0001>
- Rappaport Hovav, Malka & Beth Levin. 2008. The English dative alternation: The case for verb sensitivity. *Journal of Linguistics* 44(1). 129–167. <https://doi.org/10.1017/S0022226707004975>
- Richards, Marc. 2008. Defective Agree, case alternations, and the prominence of person. In Marc Richards & Andrej Malchukov (eds.), *Scales. Linguistische Arbeitsberichte*, Volume 86, 137–161. Leipzig: Universität Leipzig.
- Silverstein, Michael. 1986. Hierarchy of features and ergativity. In Pieter Muysken & Henk van Riemsdijk (eds.), *Features and projections*, 163–232. Dordrecht: Foris Publications Holland. <https://doi.org/10.1515/9783110871661-008>
- Tigău, Alina. 2011. *Syntax and interpretation of the direct object in Romance and Germanic languages with an emphasis on Romanian, German, Dutch and English*. Bucharest: Editura Universității din București.
- Tigău, Alina. 2014. The two-object construction in Romanian and German. In Ruxandra Cosma, Stefan Engelberg, Susan Schlotthauer, Speranța Stănescu & Gisela Zifonun (eds.), *Komplexe Argumentstrukturen. Kontrastive Untersuchungen zum Deutschen, Rumänischen und Englischen* (Konvergenz und Divergenz Series), 85–141. Berlin: de Gruyter Mouton. <https://doi.org/10.1515/9783110343229.85>
- Tigău, Alina. 2020. *Experimental insights into the syntax of Romanian ditransitives*. Berlin: de Gruyter Mouton. <https://doi.org/10.1515/9781501513657>
- Tigău, Alina & Klaus von Stechow. Manuscript. Binding properties of ditransitive constructions in Romanian.

Syntactic vs pragmatic passive

Evidence from Romanian

Andra Vasilescu

University of Bucharest / 'Iorgu Iordan – Alexandru Rosetti' Institute
of Linguistics

This chapter discusses the relationship between the two passive constructions in Romanian (*se*-passive and *be*-passive) in a twofold comparison: old Romanian vs present-day Romanian, and standard vs substandard present-day Romanian. Stipulating the distinction 'syntactic passive' vs 'pragmatic passive', I argue that the evolution of the two competing structures illustrates a typical case of convergence and I invoke the situation in present-day subdialectal Romanian as an additional proof. In old Romanian, *se*-structures functioned as syntactic passives under the Slav(on)ic influence, while *be*-structures had a temporal or copular meaning, pragmatically implying the passive meaning. In present-day standard Romanian, *be*-structures are grammaticalized for the passive meaning following the influence of Western Romance languages, while *se*-structures have acquired an impersonal-presentative meaning, continuing to pragmatically imply the passive meaning. The situation in present-day dialectal spoken varieties, more conservative compared to standard Romanian, reflects the grammaticalization/degrammaticalization processes that affected the evolution of the two structures: while *se*-structures followed the long-term evolution in standard Romanian and lost their grammaticalized passive function acquiring an impersonal-presentative meaning, the more recent grammaticalization process of *be*-structures as passives has not fully extended to subdialectal varieties. Hence, active constructions are the preferred option across subdialects, which currently lack a grammaticalized passive construction, as the passive meaning is pragmatically implied both by *se*-structures and *be*-structures.

Keywords: Romanian, *se*-passive/*be*-passive, syntactic passive/pragmatic passive

1. Introduction

Passive constructions have been analyzed from various perspectives: morpho-syntactic and semantic (Shibatani 1988; Haspelmath 1990; Klaiman 1991; Fox et al. 1995; Sohn 2000; Reinhart 2002; Blevins 2003; Reinhart & Siloni 2005; Malchukov & Siewierska 2011, among others), pragmatic (Givón 1994, for example), and cognitive (Kallmeyer 2002, for example). In many languages (like English, German, Russian, Romance languages, etc.), they are currently related to active transitive constructions: the direct object in the active structure becomes the passive subject, and the active subject becomes the agentive phrase, which possibly triggers morphological changes. Passive sentences construe the event from the patient's perspective and are pragmatically exploited to flout quantity-quality maxims, as hedging devices, or as politeness strategies. This is the case for Romanian as well.

At a glance, it appears that languages can be classified according to the number of passive structures they allow: languages with one passive structure (like English, which displays solely *be*-passives, or Slavic languages, which display solely *se*-passives), and languages which allow two passive structures (like Romance languages, which display both *se*-passives and *be*-passives). Hence a typological 'Passive Parameter' might be stipulated to account for the number of passive constructions across languages ('one-passive languages' vs 'two-passive languages'), as well as for the realization of the passive constructions in most, if not all, Indo-European languages: 'be-passive languages'/'se-passive languages'/'be- and se-passive languages'.

The relationship between the two Romanian passive structures has been a controversial issue among linguists: while Romanian linguists have considered both patterns to be Latin descendants that have simultaneously functioned as passives from the oldest attested texts, Posner (1996: 180–181) argues that the passive meaning was conveyed by *se*-structures in old Romanian under the Slavonic influence, while *be*-structures have become the canonical passive in modern Romanian under the influence of Western Romance languages.

Based on the latest descriptions of various stages in the evolution of Romanian, the present chapter compares *be*-passives and *se*-passives in old Romanian and in present-day standard and substandard Romanian. The conclusion is closer to Posner's approach, but a further distinction is introduced, i.e., syntactic passive vs pragmatic passive, to account for the simultaneous presence of both structures in all the stages of the language evolution.

To account for the differences between *se*-passive and *be*-passive, I propose the distinction 'syntactic passive' vs 'pragmatic passive'. Following Givón (1982), syntactic passives are defined as grammaticalized structures used to convey the

passive meaning, and pragmatic passives as structures which can be assigned the passive reading via conventional implicatures.¹

I make two claims: (i) both *be*-structures and *se*-structures were inherited in Romanian from vulgar Latin and were equally licit candidates for the passive meaning, but only one at a time functions as syntactic passive, while the other one serves a different function and acquires the passive reading via ‘conventional implicatures’; (ii) due to the asynchronous evolution of standard Romanian and its subdialectal varieties, present-day non-standard spoken varieties display solely a pragmatic passive.²

The two syntactic structures – i.e., the periphrastic passive (participle of the verb conjugated + *sum*) and the mediopassive (*se* + verb) – were both attested in late spoken Latin (Väänänen 1981: 127–130) and transmitted to all the Romance languages (Cennamo 1993), Romanian and its southern dialects included. They were both attested in the oldest preserved Romanian texts, both texts written directly in Romanian, and translations from Slavonic church literature. For details on passives in old Romanian, see Chivu (2000: 53–54); Timotin (2000a, 2000b, 2002); Frâncu (2009); Zafiu (2012/2015); Stan (2013); Cornilescu & Nicolae (2014); Dragomirescu (2016: 259–270); Vasilescu (2016: 188–202); Vasilescu (2017a). For an overview of present-day Romanian passives, see Pană Dindelegan & Maiden (2013). For a corpus analysis of dialectal spoken Romanian, see Vasilescu (2017b).

My first claim (see § 2 and § 3 below) is that the evolution of the two Romanian passive constructions is a typical case of convergence, as defined by Hickey (2010: 19).³ Although both structures inherited from Latin were available options in the system of Romanian and equally licit candidates to convey the passive meaning, only one of them functioned as grammaticalized passive at a given stage in the language evolution. Specifically, in old Romanian, due to the contact with Slavonic and Slavic languages which used the reflexive as a passive marker, *se*-structures were activated to canonically convey the passive meaning. They were extremely frequent and by the end of the 18th century they generalized across registers and subdialectal varieties. During this period, *be*-structures were also in use, but they preserved

1. Givón (1982) distinguishes between the syntactic mode and the pragmatic mode of encoding grammatical information across languages.

2. I use the term conventional implicature as defined by Grice (1975).

3. “A feature in a language X has an internal source, i.e., there is a systemic motivation for the feature within language X, and the feature is present in a further language Y with which X is in contact. Both internal and external sources ‘converge’ to produce the same result” (Hickey 2010: 19).

the temporal or attributive meaning they had in late Latin, and, due to the [+anteriority, +perfective] features of the participle, they conventionally implicated the passive meaning of transitive verbs.⁴ After a long-term orientation towards Eastern cultures, from the end of the 18th century, due to its orientation towards Western Romance languages and cultures, Romanian has undergone a ‘Relatinization’ process. Consequently, under the ever stronger influence of (written) French (and Italian) at the end of the 19th century, the *be*-structure has been activated and gradually grammaticalized to express the passive meaning, first in scientific texts, journalism, and translations, and then in standard spoken Romanian. In the meanwhile, after the end of the 18th century, the *se*-structure has gradually acquired an impersonal-presentative meaning, both for transitive and intransitive verbs. It continued to conventionally imply the passive meaning via the *de/de către*-phrase which parenthetically adjoins to the sentence node, as the speaker attempts to bring further clarifications about the underspecified actor involved in the event.

My second claim (see § 4) is that the present-day usage of *se*- and *be*-structures in conservative subdialectal spoken Romanian shows an intermediate stage in the parallel evolution of the two structures. On the one hand, *se*-structures in subdialectal Romanian have acquired the presentative meaning due to an evolution synchronized with standard Romanian; on the other hand, *be*-structures have preserved the status they had in old Romanian (*be* predicate + adjectival participle with stative-presentative or temporal-resultative meaning), as the passive meaning, newly acquired in standard Romanian by the end of the 19th century, has not been assimilated in conservative subdialects. Consequently, there is no dedicated syntactic passive structure in subdialectal Romanian, but the passive meaning is pragmatically implicated via the parenthetical agentive *de către*-phrase. The absence of a syntactic passive in the presence of a pragmatic passive solely appears to be consistent with the preference for the active voice in substandard and colloquial usages.

4. According to Väänänen (1981: 129) the source of the *be*-passive in late Latin is the ambiguity of structures like *domus clausa est*, which reads either as a past event (*the house had been locked in the past*) or as a present state viewed as the effect of a past event (*the house has been locked*). In the latter structure, the participle is considered to have an attributive reading, i.e., the participle attributes the present state of events to the grammatical subject, exactly like an adjective in the structure *domus parva est* (“the house is small”).

2. The evolution of *se*-structures

Se-structures are formed with the reflexive marker *se* + the lexical verb inflected for mood, tense, person, and number, agreeing with the postnominal, as shown in (1) below.

- (1) *S-au cumpărat toate cărțile.*
 SE⁵=AUX.PRF.3PL buy.PTCP all.F.PL book.F.PL.DEF.NOM=ACC
 “All the books have been/were bought.”⁵⁶

The syntactic-semantic features of *se*-structures in old Romanian differ from those in present-day Romanian. In old Romanian *se*-structures had a 3-person paradigm, allowed [+/-human], [+/-specific] subjects with an ‘entity’ or ‘property’ reading, and hosted both perfective and imperfective verbs, while the agentive phrase was frequently lexicalized (detailed in Berea Găgeanu 1966; Zafiu 2012/2015; Cornilescu & Nicolae 2014), as illustrated in (2) below.

- (2) a. *se va aduce și altă față*
 SE AUX.FUT.3SG bring.INF and other face.F.SG.NOM=ACC
 “another person will be brought too” (Prav.1581: 253^{r-v})
- b. *de tine a mă boteza*
 by you.ACC A_{INF} SE.1SG baptize.INF
 “to be baptized by you” (CC².1581: 505)
- c. *Să se știe venitul*
 SĂ_{SBJ} SE know.SBJ.3SG income.NEUT.SG.DEF.NOM=ACC
 “the income to be known” (DÎ.1599–1600, XXIX)
- d. *Decheval biruindu-se, au fugit*
 Decebal.NOM defeat.GER=SE AUX.PRF.3SG run-away.PTCP
 “Being defeated, Decebal ran away” (CIst.1700–1750: 19^r)
- e. *Leșii să mântuie de cătră Vasilie-vodă*
 Pole.M.PL.DEF.NOM SE save.PRES.3PL by Vasilie-vodă.NOM=ACC
 “The Poles are saved by Vasilie-vodă.” (NL.~1750–1766: 18)

5. Here I use SE generically, as the symbol for the Romanian reflexive clitic with a passive/impersonal reading. Nevertheless, in Romanian *se* displays several values: inherent reflexive, middle reflexive, anticausative, passive, and impersonal.

6. Romanian does not have a dedicated present perfect form, which, in certain contexts, is syncretic with the compound past (Fr. *passé composé*).

In present-day Romanian *se*-structures have a defective paradigm confined to the 3rd person and allow a [–human] or [+human, –specific] postverbal noun with a property reading, exclusively (3).⁷

- (3) a. *S-au anunțat rezultatele.*
 SE=AUX.PRF.3PL announce.PTCP result.NEUT.PL.DEF.NOM=ACC
 “The results have been announced.”
- b. *S-au cumpărat cărți pentru bibliotecă.*
 SE=AUX.PRF.3PL buy.PTCP book.F.PL.NOM=ACC for library
 “Books have been bought for (the) library.”
- c. *Se caută ingineri.*
 SE look-for.3SG engineer.M.PL.NOM=ACC
 “Engineers needed.”

The comparison indicates that *se*-structures had a genuine passive meaning in old Romanian but developed an impersonal-presentative meaning in present-day Romanian: the structure is currently used to present a state of affairs, backgrounding the agent-patient relationship.⁸ Moreover, while passive *se*-constructions in old Romanian are categorical double judgments with a topic-comment informational structure, impersonal-presentative *se*-constructions in present-day Romanian are simple,thetic judgments which lack the topic-comment informational structure.⁹ Hence, both transitive (3) and intransitive verbs (4) occur in presentative

7. Fossilized forms for the 1st and 2nd person are exceptionally preserved:

- a. *Nu te vezi bine de acolo.*
 Not CL.ACC.2SG see.PRES.2SG well from there
 “People cannot see you well from there (if you sit there).”
- b. *Nu te auzi bine de acolo*
 not CL.ACC.2SG hear.PRES.2SG well from there.
 “People cannot hear you well from there (if you sit there)”
- c. *Te cunoști când ești obosit.*
 you remark.PRES.2SG when be.PRS.2SG tired
 “People (easily) notice when you are tired.”
- d. *M-am ales președinte*
 CL.ACC.1SG=AUX.PRF.1SG elect.PTCP president
 “I was elected president.”

8. The terms ‘agent’ and ‘patient’ are generically used as cover terms for the semantic roles which the subject and the direct object, respectively, can have in the active counterparts of passive structures.

9. Kuroda (1972) and Sasse (1987) distinguished between categorical judgments and thetic judgments. Categorical judgments are double judgments which combine two cognitive acts, i.e., the first one, an act of recognizing the existence of a subject, and the second one, an act

se-structures and share the same discursive meaning: the asserted state of affairs is foregrounded, and the actor(s) involved in the event is/are backgrounded.

- (4) *Se muncește mult.*
SEWORK.PRS.3SG hard
“They (generic) work hard.”

Formerly a marker of the passive meaning in old Romanian, *se* has become a marker of the presentative-stative meaning in present-day Romanian.

As a compensatory mechanism replicating the former agentive phrase, the *de/de către*-phrase can be parenthetically adjoined to presentative-stative *se*-structures of transitive verbs (5a) and sometimes, by analogy, of some intransitive verbs too (5b).

- (5) a. *Se vor aduce probe de către*
SE AUX.FUT.3PL bring.INF proof.F.PL.NOM=ACC by
autoritățile luni.
authority.F.PL.ACC Monday
“Proofs will be brought by (the) authorities on Monday.”
b. *Se muncește mult de toată lumea.*
SEWORK.PRS.3SG hard by everybody
“Everybody works hard.”

Unlike in old Romanian *se*-passives, where the agent phrase occupied a low position under the VP and the passive subject was an IP specifier, in present-day *se*-presentatives the *de/de către*-phrase adjoins to the presentative sentence node as a parenthetical clarification of the actor(s) responsible for the state of affairs presented in the sentence.¹⁰ The relationship between the transitive verb and the demoted actor conventionally implies a passive reading within the presentative frame: “[the state of affairs presented in the sentence] has been generated by X”.

To sum up, the evolution of *se*-structures was *passive > impersonal-stative-presentative*. In old Romanian, *se*-constructions were genuine passives with a 3-person paradigm and a topic–comment informational structure, which continued the mediopassive pattern in late spoken Latin, whose passive meaning was enhanced by Slavonic/Slavic models. After the second half of the 18th century *se*-constructions have gradually changed to 1-person paradigm stative-presentatives with a thetic

of accepting/rejecting the content expressed relative to the subject. Categorical judgments are mapped as topic (subject) – comment (the predicate and its arguments). Thetic judgments are simple judgments which represent the content of a judgment and do not have a topic – comment structure.

10. Free word order in Romanian explains linearizations like 5(a), where the *de către*-phrase precedes the adverbial.

judgments reading. They can conventionally implicate the passive meaning via the sentence external *de/de către*-phrase, which discursively retrieves the entity responsible for the state of affairs described in the propositional content of the sentence. Briefly, *se*-structures functioned as syntactic passives in old Romanian, but function as pragmatic passives in present-day Romanian.

3. The evolution of *be*-structures

In Romanian, *be*-structures are formed with the auxiliary *be* inflected for mood, tense, person, and number + the participle of the lexical verb agreeing in gender and number with the syntactic subject, as shown in (6).

- (6) a. *Toate cărțile au fost cumpărate*
 all book.F.PL.DEF.NOM AUX.PRF.3PL be.PTCP buy.PTCP.F.PL
de student.
 by student.M.PL.ACC
 “All the books have been bought by (the) student.”

In certain contexts, the *be* + participle structure (7a) is ambiguous between a copular reading (7b) and a passive reading (7c).

- (7) a. *Ion a fost agitat.*
 Ion AUX.PRF.3SG be.PTCP agitate.PTCP.M.SG
 “Ion was agitated.”
- b. *Ion a fost agitat și a plâns*
 Ion AUX.PRF.3SG be.PTCP agitate.PTCP.M.SG and AUX.PRF.3SG cry.PTCP
toată ziua.
 all.F.SG day.F.SG.DEF
 “Ion was agitated and cried all day long.”
- c. *Ion a fost agitat de frații*
 Ion AUX.PRF.3SG be.PTCP agitate.PTCP.M.SG by brother.M.PL.DEF.ACC
mai mari.
 more old.M.PL
 “Ion was agitated by his elder brothers.”

Romanian also displays active participles of both transitive (8a) and intransitive verbs (8b).

- (8) a. *Ioana e mâncată.*
 Ioana be.PRF.3SG eat.PTCP.F.SG
 “Ioana has eaten.”

- b. *Sunt mulți oameni rămași acasă.*
 be.PRES.3PL many.M.PL people.M.PL remain.PTCP.M.PL home
 “Many people have remained home.”

The syncretisms displayed by *be* + participle structures in present-day Romanian replicate the systemic characteristics of old Romanian, which continued the situation in late spoken Latin described by Väänänen (1981) (see note 4). In old Romanian the *be* + participle structures were not grammaticalized to express the passive voice, but primarily displayed a stative adjectival or a perfective resultative reading.¹¹ The loose connection of the components of the structure supports this interpretation: (i) the *be* – participle order was free (9a) and (ii) allowed various (extensive) insertions (9b–c); (iii) the ellipsis of the verb, echoing the behaviour of Slav(on)ic copular structures, often positioned the participle adjacent to a noun (9d); (iv) the PP in the structure was headed by various lexical prepositions that marked other semantic cases than the agent (i.e., locative/source or instrument) (9e–f); (v) the *de/de către*-phrase used to be selected by nominalized participles as well (9g).

- (9) a. *întrebatu fu Isus de farisei*
 ask.PTCP.M.SG be.PRF.3SG Jesus by Pharisees.ACC
 “Jesus was asked by the Pharisees” (CT.1560–1561: 159^v)
- b. *este mănăstirea Solca obârșită de dânsul*
 be.3SG monastery.DEF.F.SG Solca erect.PTCP.F.SG by him
 “the monastery Solca is founded by him” (CLM.1700–1750: 86)
- c. *ce iaste de dânsul zidită*
 which be.PRS.3SG by him build.PTCP.F.SG
 “which is built by him” (ULM.~1725: 110^r)
- d. *lucru negândit și niciodată crezut*
 thing.M.SG NEG-think.PTCP.M.SG and never believe.PTCP.M.SG
 “a thing never thought of and never believed” (GIst.~1750: 124)
- e. *putearea ... ne e dată noao de la Domnul nostru*
 power.DEF.F.SG CL.DAT.1PL. be.PRS.3SG give.PTCP.F.SG. US.DAT from
 at Lord.DEF.ACC our
 “power is given to us from/by our Lord” (Ev.1642: 272)
- f. *den proroci și den apostoli propoveduită*
 from prophet.M.PL and from apostle.M.PL spread.PTCP.F.SG
 “spread by the prophets and the apostles” (NT.1648: 1^v)

11. The same diachronic evolutions and structural resemblances have been noticed for English too (Givón 1994: 4–5): *It was broken by someone/It was broken/It has been broken* vs *It was big*.

- g. *blestemaților de tatăl meu*
 curse.PTCP.M.PL.DEF.VOC by father.DEF.ACC my.POSS.M.SG
 “you, who were cursed by my father” (CS_{XIV}.1609–1618: 111^v)

At the same time, the participle had strong adjectival features: (i) active participles of various verbs were frequent (10a,c); (ii) participles frequently combined with degree markers (10c), and (iii) they selected the negative adjectival prefix *ne-* (instead of the adverbial *nu* which marks the negative of finite forms) (10b).

- (10) a. *a-i slobozi pre ei nemâncați*
 a.INF=CL.ACC.3PL liberate.INF DOM them.M NEG-eat.PTCP.M.PL
 “to set them free hungry” (NT.1648: 20^v)
 b. *oameni neînvățați* (NL.~1750–1766: 3)
 people.M.PL NEG-learn.PTCP.M.PL
 “uneducated people”
 c. *cei ce mai știuți și mai învățați*
 those.M.PL who more know.PTCP.M.PL and more learn.PTCP.M.PL
sânt (BIU.1763: V^r)
 be.PRS.3PL
 “the more knowledgeable and educated ones”

Moreover, the structure was not stable: it allowed the substitution of the *be*-verb with verbs that conveyed a modal or aspectual reading (Vasilescu 2017a; cf. Dragomirescu & Nicolae 2012), as exemplified under (11) with the verb *veni* (“come”).

- (11) *unii vinu închiși în temniță, alții*
 some.M.PL come.PRS.3PL imprison.PTCP.M.PL in gail other.M.PL
vin tăiaț (FN.1693–1704: 94)
 come.PRS.3PL cut.PTCP.M.PL
 “some of them are (to be) imprisoned, others (are to be) killed /beheaded”

A discursive argument could be added: *be*-structures, and not *se*-structures, were used in curses (Berea Găgeanu 1966; Timotin 2000a: 227–231) exactly to describe the effect in the present of a previously performed action.

Several features of *be*-structures in present-day standard Romanian indicate grammaticalization: (i) the ellipsis of the verb *be* is limited to a few cases based on coordination (12a); (ii) the *be* – participle order is fixed, and only few insertions and inversions are allowed (12b); (iii) the agent-phrase is headed solely by the dedicated prepositions *de* and *de către* (12b); (iv) the independent attributive/temporal/resultative reading of the participle faded out; (v) few active participles are still in use; and (vi) *be* – lexical verbs substitutions are exceptional.

- (12) a. *Casa a fost întâi închiriată și apoi vândută.*
 house.F.SG.DEF AUX.PRF.3SG be.PTCP first rent.PTCP.F.SG and
 then sell.PTCP.F.SG
 “The house was first rented and then sold.”
- b. *El a fost mereu apreciat de colegi.*
 he AUX.PRF.3SG be.PTCP always appreciate.PTCP.M.SG by
 colleague.PL.ACC
 “He has always been appreciated by his colleagues.”

Be-structures are especially frequent in the high registers influenced by Western models, i.e., scientific texts, journalism, and translations (documented in Pană Dindelegan 2008: 139).

To sum up, the evolution of *be*-structures was stative adjectival/perfective resultative > *passive*, as indicated by the syntactic and semantic features of the construction in old and present-day Romanian, respectively.

4. Evidence from dialectal oral varieties of Romanian

The discourse analysis of a subdialectal corpus (Vasilescu 2017a) showed that the active voice is preferred in oral interactions, while both *se*- and *be*-structures are less frequent and not spontaneously used by speakers with little instruction, but rather induced by the interviewer during the dialectal inquiry.

The interviewees most often ignore the passive structures used by the interviewer and answer with active structures (13a), they briefly echo the passive structures and then shift to active structures (13b) or, after a code-switching attempt, they abandon them (13c). The self-corrections initiated by the speakers who first use the passive structure and then resume it in the active voice (13d) show that the preferred option is the active agentive form, either with a referential or with a generic lexical/*null* subject. The discursive opposition speakers spontaneously instantiate (presentative agentless *se*-structure vs active agentive structure) confirms the analysis proposed for *se*-structures in the standard language.

- (13) a. – *Cum se face pâinea pe aici?*
 – *Cum făcea mămica înainte când eram eu mică ...* (TD: 273)
 “– How is bread made here?
 – How my mother used to make it when I was a little girl...”

- b. – *Cum se făceau clăile?*
 – *Uite, vedeți dumneavoastr-așa se făcea: puneam un snop așa jos.* (TD: 136)
 “– How were haystacks made?
 – Well, you see, that’s how they were made: we used to put a bundle on the ground.”
- c. *Iarna? s-a luat oile de la cioban, iarna oile se-ntrețin în felul următor. Le-a luat de la cioban.* (TD: 151)
 “In winter? The sheep had been taken away from the shepherds. In winter sheep are taken care of like that: they had taken them from the shepherd.”
- d. *Se dă darurile ș-acolo joacă ... invitații ... și dă daruri.* (TD: 410)
 “Presents are offered and they dance there ... the guests ... and they offer presents.”

As for *be*-passives (by far fewer than *se*-structures), they are introduced by the interviewer, but the interviewee immediately shifts to the active form (14a). There are few cases when speakers spontaneously use *be*-passive structures, and that is to quote administrative documents or the speech style of those whom they perceive to be educated speakers (14b). Other instances of *be*-structures seem to preserve the copular attributive meaning in old Romanian (14c). Aspectual structures (14d) are frequent, like in old Romanian and unlike in present-day standard Romanian.

- (14) a. – *Dar dumneata știi cum erau lucrate?*
 – *E! le tăia așa și le făcea găuri așa.* (TD: 893)
 “– But do you know how they were made?
 – Well! They used to cut them like that and make holes like that.”
- b. *ea va fi amendată cu cinci sute de lei* (TD: 526)
 “she will be fined 500 lei”
- c. *secerători dintr-alea ... e trase de boi, știți?* (TD: 398)¹²
 “harvesters of that kind ... oxen-pulled ones, you know?”
- d. *vine despărțită-n trei* (TD: 256)
 come.PRS.3SG split.PTCP.F.SG=in three
 “it is split in three”

Be-passives are little used in oral subdialectal varieties, which suggests that they are mainly an option of the literary use.

12. In this example, *e trase de boi* (lit. “they are pulled by oxen”) rather means “they are oxen-pulled” and seems to contrast with “they are mechanical”, and not with an active counterpart “the oxen pull them”. In terms of constituency, *e trase de boi* has the structure [they [are [pulled by oxen]]], not [they [[are pulled] by oxen]].

To sum up, in oral interactions speakers with little instruction rather prefer active structures than non-active ones. *Se*-structures enter the discursive opposition active/impersonal-presentative as in standard language, while the few *be*-structures found in the corpus either imitate the literary use or have an attributive or resultative reading. It appears that dialectally spoken Romanian does not have a grammaticalized passive voice, but derives passive meaning via conventional implicatures, as showed in § 2 and § 3 above.

4. Conclusions

This chapter challenges current opinions on the existence of two parallel passive structures in Romanian. The claim is that Romanian displays one syntactic passive and one pragmatic passive in each of the two stages of evolution under research (§ 1). On the one hand, the comparison between old and present-day standard Romanian showed that *se*-structures display different syntactic features in the two stages, indicating their evolution from a genuine passive meaning to an impersonal-presentative meaning, i.e., from a syntactic passive to a pragmatic passive (§ 2). On the other hand, the comparison between old and present-day standard Romanian showed that *be*-structures also display different syntactic features in the two stages, indicating an evolution from a temporal-aspectual-attributive/stative-resultative meaning to a genuine passive meaning, i.e., from a pragmatically implicated passive to a syntactic passive (§ 3). The evolution of the two structures appears to be a typical case of convergence. While both structures were inherited from Latin, they developed different meanings in use depending on the dominant external influence: the influence of Slavonic/Slavic languages before the end of the 18th century, and the influence of Western (Romance) languages after that date. The analysis of the usage of *se*-structures and *be*-structures in conservative subdialectal varieties of Romanian nowadays (§ 4) shows the prevalence of active structures over passive ones, the personal/impersonal opposition instantiated by active vs *se*-structures, and very few *be*-structures, which actually imitate educated speech style. These findings were interpreted as direct evidence for the evolution of the two structures in standard language. On the one hand, the evolution of *se*-structures in standard and substandard Romanian was synchronized due to the long time-span (until the end of the 18th century) and to the lax character of standard language norms during the period. On the other hand, the more recent evolution of *be*-structures under the Western influence (after the end of the 18th century) is not synchronized due to the conservative character of subdialects and to a more constraining norm of the literary usage.

The present account of the double Romanian passive is hybrid in nature. It partly confirms Posner's analysis but refines it with the distinction syntactic passive/pragmatic passive, which allows to partly preserve the interpretation proposed by Romanian linguists, who noticed the circulation of both structures over time.

Acknowledgements

Special thanks to Meredith Walker, retired lecturer from Clemson University, South Carolina, USA, who accepted to do the language review of my article.

Corpus

- BIU.1763 – Vlad Boțulescu de Mălăiești, *Scrieri III. Istoria universală. Asia*, ed. Emanuela Timotin, Andrei Timotin, Bucharest: Univers Enciclopedic Gold, 2003, 101–195.
- CC².1581 – Coresi, *Evanghelie cu învățătură*. Ed.: Diaconul Coresi, *Carte cu învățătură (1581), Volume 1, Textul*, ed. Sextil Pușcariu, Al. Procopovici, Bucharest: Socec, 1914.
- CIst.1700–1750 – Constantin Cantacuzino, *Istoria Țării Românești*. Ed.: *Istoria Țării Rumânești atribuită stolnicului Constantin Cantacuzino*, ed. Otilia Dragomir, Bucharest: Editura Academiei Române, 2006, 145–202.
- CLM.1700–1750 – Miron Costin, *Letopisețul Țării Moldovei*. Ed.: M. Costin, *Opere*, ed. P. P. Panaiteșcu, Bucharest: Editura de Stat pentru Literatură și Artă, 1958, 41–201.
- CS_{xiv}.1609–1618 – *Codex Sturdzanus. Tâlcul evangheliei de la judecată (Zise Domnul)*. Ed. Gh. Chivu, Bucharest: Editura Academiei Române.
- CT.1560–1561 – Coresi, *Tetraevanghel*. Ed.: *Tetraevanghelul tipărit de Coresi*. Brașov 1560–1561, comparat cu Evangheliarul lui Radu de la Mănăcești 1574, ed. Florica Dimitrescu, Bucharest: Editura Academiei, 1963.
- DÎ – *Documente și însemnări românești din secolul al XVI-lea*, text stabilit și indice de Gh. Chivu, Magdalena Georgescu, Magdalena Ioniță, Alexandru Mareș, Alexandra Roman-Moraru, Bucharest: Editura Academiei Române, 1979.
- Ev.1642 – *Evanghelie învățătoare*. Ed. Alin-Mihai Gherman, Bucharest: Editura Academiei Române, 2011, 153–480.
- FN.1693–1704 – *Foietul Novel. Calendarul lui Constantin Vodă Brâncoveanu (1693–1704)*, ed. Emil Vârtosu, Bucharest: Monitorul Oficial și Imprimeriile Statului, 1942, 4–194.
- GIST.~1750 – Radu Greceanu, *Începătura istoriei vieții luminatului și preacreștinului Domnului Țării Românești*. Ed. *Cronicari munteni*, Volume 2, ed. Mihail Gregorian, Bucharest: Editura pentru Literatură, 1961, 5–272.
- NL.~1750–1766 – Ion Neculce, *Letopisețul*. Ed.: Ion Neculce, *Letopisețul Țării Moldovei și O samă de cuvinte*, ed. Iorgu Iordan, Bucharest: Editura de Stat pentru Literatură și Artă, ed. a II-a, 1959, 31–388.
- NT.1648 – *Noul Testament*. Ed. Alba Iulia: Reîntregirea, 1998.
- Prav.1581 – *Pravila ritorului Lucaci*. Ed. I. Rizescu, Bucharest: Editura Academiei, 1971, 161–183.

- TD – *Texte dialectale. Muntenia*. III, de Costin Bratu, Galina Ghiculete, Maria Marin, Bogdan Marinescu, Victorela Neagoie, Ruxandra Pană, Marilena Tiugan, Magdalena Vulpe, 1987 [1976]. Bucharest: [Editura Academiei].
- ULM.~1725 – Grigore Ureche, *Letopiseșul Țării Moldovei*. Ed. P. P. Panaitescu, Bucharest: Editura de Stat pentru Literatură și Artă, 1955, 57–210.

References

- Berea Găgeanu, Elena. 1966. Observații asupra diazezei pasive în limba română. *Studii și cercetări lingvistice* 17(4). 567–578.
- Blevins, James P. 2003. Passives and impersonals. *Journal of Linguistics* 39. 473–520. <https://doi.org/10.1017/S0022226703002081>
- Cennamo, Michaela. 1993. *The reanalysis of reflexives: A diachronic perspective*. Naples: Liguori.
- Chivu, Gheorghe. 2000. *Limba română de la primele texte până la sfârșitul secolului al XVIII-lea. Variantele stilistice*. Bucharest: Editura Univers Enciclopedic.
- Cornilescu, Alexandra & Alexandru Nicolae. 2014. The grammaticalization of a constraint on passive reflexive constructions in Romanian. In Gabriela Pană Dindelegan, Adina Dragomirescu, Irina Nicula, Alexandru Nicolae & Louise Esher (eds.), *Diachronic variation in Romanian*, 309–362. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Dragomirescu, Adina. 2016. The past participle and the participial clause. In Gabriela Pană Dindelegan & Martin Maiden (eds.), *The syntax of old Romanian*, 259–270. Oxford: Oxford University Press.
- Dragomirescu, Adina & Alexandru Nicolae. 2012. Pasivul cu *a veni* – traducere din italiană sau inovație românească?. *Limba română* 61(1). 73–81.
- Fox, Barbara & Paul J. Hopper (eds.). 1995. *Voice: Form and function*. Amsterdam: John Benjamins.
- Frâncu, Constantin. 2009. *Gramatica limbii române vechi (1521–1780)*. Iași: Casa Editorială ‘Demiurg’.
- Givón, Talmy. 1982. Transitivity, topicality, and the Ute impersonal passive. In Paul J. Hopper & Sandra A. Thompson (eds.), *Syntax and semantics 15: Studies in transitivity*, 143–160. New York: Academic Press. https://doi.org/10.1163/9789004368903_009
- Givón, Talmy. 1994. The pragmatics of de-transitive voice: Functional and typological aspects of inversion. In Talmy Givón (ed.), *Voice and inversion*, 3–44. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.28.03giv>
- Grice, Herbert Paul. 1975. Logic and conversation. In Peter Cole & Jerry Morgan (eds.), *Syntax and semantics 3: Speech acts*, 41–58. London: Academic Press.
- Haspelmath, Martin. 1990. The grammaticalization of passive morphology. *Study of Language* 14(1). 25–72. <https://doi.org/10.1075/sl.14.1.03has>
- Hickey, Raymond. 2010. Language contact: Reconsideration and reassessment. In Raymond Hickey (ed.), *The handbook of language contact*, 1–28. Oxford: Wiley-Blackwell. <https://doi.org/10.1002/9781444318159.ch>
- Kallmeyer, Werner. 2002. Verbal practices of perspective grounding. In Carl Friedrich Graumann & Werner Kallmeyer (eds.), *Perspective and perspectivation in discourse*, 113–142. Amsterdam: John Benjamins. <https://doi.org/10.1075/hcp.9.09kal>

- Klaiman, Miriam H. 1991. *Grammatical voice*. Cambridge: Cambridge University Press.
- Kuroda, Sige-Yuki. 1972. The categorical and the thetic judgment. *Foundations of Language* 9. 153–185.
- Malchukov, Andrej & Anna Siewierska (eds.). 2011. *Impersonal constructions*. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.124>
- Pană Dindelegan, Gabriela. 2008. Construcții pasive și construcții impersonale. In Valeria Guțu Romalo (ed.), *Gramatica limbii române*, Volume 2, 133–147. Bucharest: Editura Academiei Române.
- Pană Dindelegan, Gabriela & Martin Maiden (eds.). 2013. *The grammar of Romanian*. Oxford: Oxford University Press.
- Posner, Rebecca. 1996. *The Romance languages*. Cambridge: Cambridge University Press.
- Reinhart, Tanya. 2002. The theta system: An overview. *Theoretical Linguistics* 28. 229–290.
- Reinhart, Tanya & Tal Siloni. 2005. The lexicon-syntax parameter: Reflexivization and other arity operations. *Linguistic Inquiry* 36(3). 389–436. <https://doi.org/10.1162/0024389054396881>
- Sasse, Hans-Jürgen. 1987. The thetic/categorical distinction revisited. *Linguistics* 25. 511–580. <https://doi.org/10.1515/ling.1987.25.3.511>
- Shibatani, Masayoshi. 1988. *Passive and voice*. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.16>
- Sohn, Joong-Sun. 2000. Passivity and impersonality of the Romance reflexive clause. *Linguistics* 8(1). 333–352.
- Stan, Camelia. 2013. *O sintaxă diacronică a limbii române vechi*. Bucharest: Editura Universității din București.
- Timotin, Emanuela. 2000a. Pasivul românesc în epoca veche. Aplicație asupra textelor originale din secolul al XVI-lea. *Studii și cercetări lingvistice* 51. 225–231.
- Timotin, Emanuela. 2000b. Modalități de exprimare a pasivului în dialectele românești sud-dunărene. *Studii și cercetări lingvistice* 51. 481–490.
- Timotin, Emanuela. 2002. Originea pasivului românesc. *Studii și cercetări lingvistice* 53. 95–109.
- Väänänen, Veikko. 1981. *Introduction au Latin vulgaire*. Troisième Edition Revue et augmentée. Paris: Editions Klincksieck.
- Vasilescu, Andra. 2016. Passive Constructions. In Gabriela Pană Dindelegan (ed.), *The syntax of old Romanian*, 188–202. Oxford: Oxford University Press.
- Vasilescu, Andra. 2017a. Diateză și perspectivă. In Adina Dragomirescu, Alexandru Nicolae, Camelia Stan & Rodica Zafiu (eds.), *Sintaxa ca mod de a fi. Omagiu Gabrielei Pană Dindelegan, la aniversare*, 461–479. Bucharest: Editura Universității din București.
- Vasilescu, Andra. 2017b. Pasivul canonic vs pasivul reflexiv în româna nonstandard. *Studii și cercetări lingvistice* 68(2). 187–217.
- Zafiu, Rodica. 2012/2015. Sintaxa. In Gheorghe Chivu, Gabriela Pană Dindelegan, Adina Dragomirescu, Isabela Nedelcu & Irina Nicula (eds.), *Studii de istorie a limbii române. Morfosintaxa limbii literare în secolul al XIX-lea și al XX-lea*, 322–323. Bucharest: Editura Academiei Române.

Index

A

- a* 94, 163
- A-Chain 10
 - ~ Deficit Hypothesis (ACDH) 10, 11
- aboutness 87, 112
- Abruzzese 207
- accent
 - lexical 135, 136
 - nuclear 135, 136
 - prenuclear 147 (see *stress*)
- acquisition 9–25
- AD 159
- adjunct
 - ~ extraction 263, 264, 265, 266, 267, 269, 273, 274, 275
- adverb
 - agent-oriented 14
 - floating 44, 46
 - presuppositional 43
- adverbial clause 121–125
- affectedness 107
- Aghem 142
- Agnone 225, 229, 231
- Agree 279, 285
- agreement
 - overt object ~ 185
 - participial ~ 17
- alors* 291
- Amandola 225, 229, 231
- animacy 182, 183, 185, 186, 338
 - ~ Hierarchy (AH) 92, 94, 95, 96, 97, 101, 105
- applicative 99, 102, 103
- argument realization 139
- Ariellese 207, 228
- aspect 196, 303
- atelic 246, 247, 248
- atransitive 259
- aucun* 34, 35, 37

autosegmental-metrical model 138

auxiliary 13, 24, 141

- ~ selection 14, 16

B

- Balkan
 - ~ language 163, 164
 - ~ Sprachbund 157
- bilingualism 163
- Bitetto 225, 227, 229, 230
- by*-phrase 13, 17, 24, 360

C

- ca* 161, 165, 219, 220, 222
- Calabrese 157–167
- Campobasso 213
- Cantonese 196, 200
- Capracotta 225, 226, 229, 230
- Case Resistance Principle (CRP) 222, 223
- Castelvecchio Subequo 225, 226, 229, 231
- Catalan 9, 15, 16, 24, 92, 252
- causative 240
 - ~ inchoative alternation 239–259
 - Stative-~ (SC) 241, 242, 243, 245, 246, 247, 248, 250, 251, 254, 255, 256, 257, 258, 259
- cã* 165
- che* 85, 161, 166
- chi* 165
- Chinese 59
- clause structure 208, 211
- cleft-construction 193, 200
- clitic doubling (CD) 78, 82, 85, 88, 91, 92, 97, 100, 102, 103, 108, 188, 190, 335, 337, 338, 340, 341, 343, 346, 348, 359, 351

- conditional clause 122–124
 - connectivity effects 50, 51
 - complementizer 158, 219
 - ~ trace effect 271
 - coordination 51, 60
 - asymmetric ~ 63
 - ~ of Likes Constraint 51
 - ~ of Unlikes 52
 - conative 298, 304, 307, 308, 309
 - construction 60
 - constructionalist approach 242, 256
 - contact
 - language ~ 163, 200, 201
 - control 125, 131, 132, 159, 245
 - exhaustive ~ 221
 - partial ~ 221
 - non-obligatory ~ 221, 222, 223
 - obligatory ~ 221, 222, 223
 - convergence 357, 359, 369
 - copula 278, 279, 292
 - coreference 219
 - Cosentino 205–232
 - creole 193, 196, 201
 - cu* 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167
 - cyclicity without containment 311–333
- ## D
- da* 159
 - dative
 - ~ alternation 91–108
 - ~ clitic 92, 175, 335–353
 - double ~ 91
 - Goal/Recipient ~ 91, 96, 103, 108
 - inflectional ~ 91, 93, 94, 95, 96, 100, 101, 108, 305, 345, 346, 347
 - multiple ~ 103, 104

- possessive ~ 91, 103, 104
 prepositional ~ 91, 92, 93,
 94, 95, 96, 97, 105, 106,
 108, 171, 298, 304, 305, 306,
 308, 309, 345, 347
de 265, 284, 301, 260, 363,
 364, 365, 366
de c atre 360, 363, 364, 365, 366
de que 263–275
 DE 159
 declarative 142
 decreolization 196
 Default Linking Rule 255
 Definiteness Hierarchy 95,
 97, 105
 degrammaticalization 357
 Differential Object Marking
 (DOM) 105, 171–190, 335,
 338, 340, 341, 342, 343, 346,
 347, 348, 349, 351, 352
 distance effect 263, 266, 267,
 273, 274, 275
 ditransitive 91–108, 335–353
donc 291
dont 278, 279, 286, 287, 289
 Double Object Construction
 92, 335, 336
 doubly-filled COMP 85
 Dutch 10, 195
- E**
- Edge Generalization 292
 ellipsis 49–67
 English 9, 10, 23, 50, 52, 59, 60,
 80, 92, 122, 141, 194, 195, 196,
 242, 251, 263, 265, 266, 271,
 337, 358
essere 13
est-ce que 278, 279, 280, 292
estar 13, 14, 15, 16, 23, 24
 event composition 139
 Extended Projection Principle
 111
- F**
- Featural Relativized Minimality
 73, 79, 80
 feature
 ~ disjunction 79
 ~ identity 79, 86, 87
 ~ inclusion 86, 87
- ficar* 13, 14, 15, 23, 24
 Fischer Test 130
 Ficiary role 96, 107, 108
 Finnish 114, 116, 117, 123, 124,
 125
 focalization 75
 in situ ~ 75
 focus 49, 65, 66, 73–88
 adverbial ~ 78, 81, 86
 Clitic-Doubled foci 78, 79,
 82, 83, 84
 contrastive ~ 83
 double foci 49, 66, 67
 early-~ (EF) 135, 138, 139,
 140, 141, 148, 149, 150,
 151, 152
 final-~ (FF) 136, 138, 139,
 147, 148, 150, 152
 ~ fronting 87
 mirative ~ 144
 narrow ~ 139, 216, 217
 polarity ~ 142
 predicate ~ 141
 Uniqueness of ~ 74
 verum ~ 135–153
 wide ~ 216, 217
 fragment answer 32, 35, 36, 67
 Freezing Principle 12
 French 10, 29–46, 50, 80, 92,
 194, 277–293, 360
 Frigento 225, 229, 230
- G**
- Galician 268, 269
 Gallo Matese 229, 230
 gapping 49–67
 Functional Sentence Perspec-
 tive Principle of ~ 65
 Generalized Linear Mixed
 Model (GLMM) 19, 21, 22
 genitive 193, 197–199, 278, 285
 German 50, 141, 242, 251, 358
 Giurdignano 225, 229
 grammaticalization 165, 357
 downward ~ 157–167
 secondary ~ 165
 Gravina di Puglia 225, 229, 231
 Greek 9, 160, 163, 239, 243, 253,
 254, 255, 256, 259
 Guardiaregia 225, 226, 229, 230
- H**
- Haitian 194
 Head driven Phrase Structure
 Grammar (HPSG) 61, 64
 Head over Phrase Principle 162
(hi)na 163, 166
 Hungarian 138
- I**
- iar* 55, 60, 65, 66
 IbexFarm 56
il 287
 Immediate Cause Linking Rule
 242, 255
 impersonal *se* 249
 inchoative 240, 242, 247, 249,
 250, 253
 infinitive 164
 infinitival complementation
 157
 information structure 50, 65,
 111–132, 208
 Interface Root Restriction 113
 Interpolation 211
 interrogative
 direct ~ 277
 (see *question*)
 intonation 135, 138, 146, 153
 intonational phrasing 206
 intransitive 240, 360, 362, 364
 inversion
 clitic ~ 280
 complex ~ (CI) 277, 278,
 284, 287, 289, 290, 292
 stylistic ~ 280
 irrealis
 ~ clause 222
 ~ subordinator 157–167
 ~ subjunctive 220
 island 263, 264, 275
because-~ 264
strong ~ 277
wether-~ 263
 Italian 10, 13, 50, 52, 73, 75, 77,
 78, 79, 80, 81, 82, 86, 88, 92,
 114, 116, 117, 118, 120, 123, 124,
 125, 126, 127, 128, 129, 130, 131,
 246, 252, 360
- J**
- Japanese 50, 59

- K**
 Korean 50, 59
 Kristang 193–201
- L**
la 91, 94, 98, 100, 297, 298, 306, 308
 Language Bioprogram
 Hypothesis (LBH) 194
 language ecology 193, 201
 Latin 160, 297, 313, 339, 358
 Late ~ 162, 360
 Vulgar ~ 359
 spoken ~ 359
liaison 206
 licensing 171–190
 Likert scale 115
 lme4 56
 locality 60, 120, 126, 132, 211, 214, 222, 224, 232
- M**
 Maglie 225, 229
 Makista 193–201
 Malay 200
 mediopassive 359
 merge
 external ~ 81, 86, 88
 Late ~ Principle 162
 ~ over Merge Principle 162
 middle
 ~ alternation 251
 Minimal Link Condition 131
 MODO 160
 Molfetta 225, 226, 229
 Monteroduni 229, 230
 movement 60, 76, 159, 165, 267, 269, 270, 272, 273, 279, 286
 long distance ~ 269
 successive cyclic ~ 263, 267, 269, 270, 271, 272, 275
mu/ma/mi 157, 158, 159, 160, 161, 162, 163, 164, 165, 166, 167
- N**
ne 30, 31, 33, 34, 35, 39, 44, 46
ne-cliticization 246
 Neapolitan 171, 172, 173, 185, 186, 188, 189, 190, 206
 negation 29–46, 143, 194
 double ~ 29, 32, 33, 35, 38, 40, 41, 42, 43, 45
 Embracing ~ 30, 33
 existential ~ 37
 half-~ 33, 34, 35
 logical ~ 35
 multiple ~ 32
 sentential ~ 29, 30, 33, 35, 36, 37, 39, 44, 45, 46
 universal ~ 37
 negative
 ~ concord 29, 31, 32, 33, 35, 36, 38, 39, 40, 41, 43, 45
 ~ doubling 33
ni 34
 non-quantificational 29–46
- O**
 object
 ~ extraction 263, 264, 269
 null ~ (NO) 243, 244, 245, 246, 247, 248, 254, 258
 oblique differential marker 173
- P**
 Padula 225, 229, 231
 parallelism 49–67
 particle
 aspect ~ 196
 genitive ~ 197
 modal ~ 158
pas 30, 31, 32, 33, 34, 35, 38, 43
 passive 278
 adjectival ~ 10, 11, 12, 15, 16, 17, 24, 25
 be-~ 357, 359, 360, 364, 365, 366, 367, 368, 369
 long ~ 15, 16, 21, 22, 23
 ~ Parameter 258
 periphrastic ~ 359
 pragmatic 357–370
 se-~ 249, 357, 359, 360, 361, 362, 363, 367, 368, 369
 short 11, 15, 16, 21, 22, 23
 syntactic 357–370
 verbal 9–25
 Pàsterna-Castelpetroso 225, 226, 229, 231
pe 105, 306, 338, 342, 346
 perfect 311–333
 analytic ~ 312, 331
 ~ correspondence 319, 327, 329
 ~ participle 312
 simple ~ 312, 328, 331
 strong ~ 325, 327, 328
 weak ~ 325, 328, 329
 Person-Case Constraint 175
personne 29, 34, 36, 37, 38, 40, 43
 phasal domain 207, 212
 phase 208
 ~ Collapsing 263, 267, 268, 269, 272, 274, 275
 ~ Impenetrability Condition (PIC) 213, 221, 227, 228, 229, 230, 231, 232, 269, 270
 ~over-~ Configuration 267, 268, 272
 ~over-~ Constraint 268
 strong ~ 221
 ~ Theory 211, 227, 232, 279
 Universal ~ Requirement (UPR) 10, 11, 15, 17
 pidgin 194
 pluperfect 312, 313, 317, 318, 319, 324, 326, 331
plus 29–46
 Poggio Imperiale 225, 226, 229, 230
 Pontecorvo 225, 227, 229, 231
 Popoli 225, 229, 231
 Portuguese 193, 196, 252
 Brazilian ~ 13, 76
 European ~ 11–25, 50
pre 105
 predicate
 complex ~ 299, 302
 preposition stranding 80
 prepositional
 ~ accusative 174, 176
 ~ object contraction 335, 336
 presupposition 29–46, 74
pro-drop 49–67, 275
 consistent ~ 111, 112, 113, 119, 126, 128, 129, 131
 partial ~ 111, 112
 non-~ 112
 semi-~ 112
 Projection Principle 111
 pronominalization 159

- purposive
 ~ adjunct clause 161, 163
 ~ complementizer 162
 ~ reading 164
 Putignano 225, 226, 229
- Q**
 quantificational chain 73–88
 quantifier
 existential ~ 36
 negative ~ 36
que 34, 161
 QUE 163
 question
 polar ~ 135–153
 (see *interrogative*)
 QUID 161
 QUOD 157, 159, 161, 162, 163
 QUOMODO 157, 160, 161, 162,
 163, 165, 167
- R**
 R 56
raddoppiamento fonosintattico
 (RF) (phonosyntactic
 doubling) 205–232
 ~ triggers 208
 realis
 ~ complementizer 165
 reanalysis 162, 268
 reconstruction 51
 relative
 reduced ~ 299, 301
rien 29, 34, 36, 37, 38, 40
Roccasicura 225, 229
 Romanian 49, 50, 52, 53, 56,
 57, 59, 61, 63, 65, 66, 67, 83,
 91–108, 111–132, 135–153,
 157–167, 171, 172, 173, 178, 179,
 180, 181, 182, 183, 186, 188,
 189, 190, 252, 255, 297–309,
 311–333, 335–353, 357–370
 Russian 114, 116, 117, 123, 358
Ruvo di Puglia 225, 229
- S**
 S. Benedetto del Tronto 225,
 229, 231
 S. Giorgio del Sannio 225, 227,
 229, 230
 S. Lorenzo del Vallo 225, 229
 S. Vittore 225, 226, 229
 Salentino 157–167
 Sandhi 205, 206
sans 37
sā 200
sā 157, 158, 159, 164, 165
se 166, 243, 249 (see *passive*)
se-si cliticization 240
 Secinaro 225, 226, 229, 231
ser 13, 14, 15, 16, 23
 si(c) 160, 164, 165, 166, 167
 Sesotho 13
 Sicilian 158
 Simpler Syntax Hypothesis 61
 Slav(on)ic 357, 358, 359, 363,
 369
 Smuggling 12
 Sonnino 225, 229, 231
 Spanish 9, 15, 16, 49, 50, 52, 53,
 56, 57, 58, 59, 65, 66, 67, 92,
 171, 172, 173, 174, 177, 178, 180,
 181, 182, 183, 189, 190, 245, 251,
 252, 253, 263–275, 315, 337, 342
 stem 327
 ~ identity 311, 312, 313, 315,
 321, 327, 331
 perfect ~ 317, 326, 327, 328,
 330
 stress 311–333
 iambic ~ 329
 lexical ~ 329
 nuclear ~ 141, 150, 153
 ~ sistem 314–317
 (see *accent*)
 structural complexity 240
 subject
 ~ Condition (SC) 277, 278,
 279, 280, 281
 ~ extraction 263, 264, 266,
 267, 273, 275, 279, 291
 ~ pseudo-extraction 290
 null ~ 49, 59, 65, 67, 111–132
 (see *pro-drop*)
 overt ~ 49, 59, 65, 67
 ~ raising 10
 subjunctive 157, 158
 supine 297–309
 clausal ~ 298
 definite ~ 302, 303, 304, 306
 nominal ~ 298, 300, 303
 prepositional (bare) ~ 298,
 302, 303, 304, 306, 307
 verbal ~ 298, 300, 302
 syntactic composition 240
ši 164
- T**
 temporal clause 124–125
teng 200
to 96, 97, 163
 tone 135, 139
 topic 49, 66, 82, 129
 Aboutness-shift ~ (A-Topic)
 112, 113, 116, 117, 118, 120,
 124, 125, 126, 127, 130, 131
 ~ chain 111, 112, 113, 118, 121,
 123, 131
 ~ Chain Condition 113,
 125, 130
 contrastive ~ 65, 82
 ~ Criterion 111, 112, 113, 125,
 126, 130, 131
 G-Topic (Aboutness,
 Background) 119, 120,
 123, 130, 132
 Hanging ~ 125
 ~ selection 113
 silent ~ 111
 topicalization 74, 83
tough-construction 299, 301
 transitive 244, 279, 358, 360,
 362, 364
 transitivity
 ~ alternation 239, 258
 transitivization 241
 transparency 159
Trevigiano 73–88
Tuffillo 225, 229
- U**
 unaccusative 240, 242, 255,
 256, 278, 279
 unergative 254, 280
 Universal Freezing Hypothesis
 (UFH) 12, 15, 17, 24
 UT 157

- V
- Vastogirardi 225, 229, 231
- venire* 13
- verb
- actional ~ 9, 11, 15, 16, 21, 24
 - aspectual ~ 159
 - bridge ~ 111, 115, 116–119
 - change of state ~ 14, 241, 242, 243, 247, 251, 255
 - factive ~ 111, 115, 119–121
 - modal ~ 159
 - motion ~ 297
 - non-actional (experiencer) ~ 9, 10, 11, 12, 15, 16, 24
 - perception ~ 16, 21, 22, 23
 - psych ~ 242, 251, 256
 - transfer of possession ~ 97
- Verbicano 207
- Viticuso 225, 229, 231
- W
- Wasow's Generalization 52
- weak cross-over effects 83
- Welsh soft mutation 205
- wh-*
- ~ chain 75
 - ~ extraction 263–275
 - ~ in situ 280
- phrase 162
- projection 73–88
- question 73–88
- Y
- y* 55
- Z
- zu* 163

This volume contains a selection of 18 peer-reviewed papers presented at the 31st edition of *Going Romance*. Phenomena found in Romance languages (European Portuguese, French, Italian, Spanish, Romanian), in Romance dialects (Cosentino, Salentino, southern Calabrese, Neapolitan, and Trevigiano), and even in creoles with a Romance lexifier (Makista and Kristang) either benefit from in-depth analyses confined to one single variety, or are subjected to comparative analysis (dialect vs standard language, dialect vs different major language(s), cross-dialectal comparison, cross-Romance comparison, and even comparison of language families). Theoretical and experimental approaches complement one another, as do diachrony and synchrony. Individually and as a whole, these contributions show how the Romance languages contribute to a better understanding of issues which are relevant in the current linguistic landscape: acquisition, n-words, ellipsis phenomena, focus and polarity, ditransitive constructions, grammaticalization theory, differential object marking, language ecology, event structure, cyclicity, passives and many more.

ISBN 978 90 272 1005 0



JOHN BENJAMINS PUBLISHING COMPANY