

Romance Languages and Linguistic Theory

14

Romance
Languages
and
Linguistic
Theory 14

Selected papers from
the 46th Linguistic
Symposium on Romance
Languages (LSRL),
Stony Brook, NY

edited by
Lori Repetti
Francisco Ordóñez

John Benjamins Publishing Company

Romance Languages and Linguistic Theory 14

Romance Languages and Linguistic Theory (RLLT)

ISSN 1574-552X

The yearly events of both 'Going Romance' and 'Linguistic Symposium on Romance languages' feature research in formal linguistics of Romance languages, in the domains of syntax, morphology, phonology and semantics. Each volume brings together a peer-reviewed selection of papers that were presented at one of the meetings, aiming to provide a representation of the spread of topics at that conference, and of the variety of research carried out nowadays on Romance languages within theoretical linguistics.

For an overview of all books published in this series, please see <http://benjamins.com/catalog/rllt>

Editor

Frank Drijkoningen
Utrecht University

Volume 14

Romance Languages and Linguistic Theory 14. Selected papers from the 46th Linguistic Symposium on Romance Languages (LSRL), Stony Brook, NY
Edited by Lori Repetti and Francisco Ordóñez

Romance Languages and Linguistic Theory 14

Selected papers from the 46th Linguistic
Symposium on Romance Languages (LSRL),
Stony Brook, NY

Edited by

Lori Repetti
Francisco Ordóñez
Stony Brook University

John Benjamins Publishing Company
Amsterdam / Philadelphia



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

DOI 10.1075/rllt.14

**Cataloging-in-Publication Data available from Library of Congress:
LCCN 2018014195 (PRINT) / 2018034380 (E-BOOK)**

ISBN 978 90 272 0097 6 (HB)
ISBN 978 90 272 6389 6 (E-BOOK)

© 2018 – 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 LSRL 46 Stony Brook	1
CHAPTER 1	
Expletive negation is not expletive: Evidence from aspect in Spanish <i>Paola Céspedes</i>	5
CHAPTER 2	
Long-distance binding of French reflexive <i>soi</i> : First-person oriented logophoricity <i>Isabelle Charnavel</i>	21
CHAPTER 3	
French negative concord and discord: An experimental investigation of contextual and prosodic disambiguation <i>Viviane Déprez and Jeremy Yeaton</i>	35
CHAPTER 4	
Dimensions of variation: The inflected construction in the dialect of Delia (Caltanissetta) <i>Vincenzo Nicolò Di Caro and Giuliana Giusti</i>	53
CHAPTER 5	
<i>Ma non era rosso?</i> (But wasn't it red?): On counter-expectational questions in Italian <i>Alessandra Giorgi</i>	69
CHAPTER 6	
<i>Dime una cosa</i> : Are wh-in-situ questions different in Spanish? Evidence from intonation <i>Carolina González and Lara Reglero</i>	85

CHAPTER 7	
Parametric comparison and dialect variation: Insights from Southern Italy	103
<i>Cristina Guardiano, Dimitris Michelioudakis, Guido Cordoni, Monica Alexandrina Irimia, Nina Radkevich and Ioanna Sitaridou</i>	
CHAPTER 8	
Morphological doublets in Brazilian Portuguese <i>wh</i> -constructions	135
<i>Mary A. Kato</i>	
CHAPTER 9	
Clitic doubling, person and agreement in French hyper-complex inversion	153
<i>Richard S. Kayne</i>	
CHAPTER 10	
Licensing conditions on null generic subjects in Spanish	185
<i>Matthew L. Maddox</i>	
CHAPTER 11	
Bridging and dislocation in Catalan	201
<i>Laia Mayol and Xavier Villalba</i>	
CHAPTER 12	
Dependent numerals and dependent existentials in Romanian	215
<i>Mara Panaitescu</i>	
CHAPTER 13	
Stressed enclitics are not weak pronouns: A plea for allomorphy	231
<i>Diego Pescarini</i>	
CHAPTER 14	
Causativization of verbs of directed motion in Romance languages	245
<i>Anna Pineda</i>	
CHAPTER 15	
Latin denominal deponents: A syntactic analysis	263
<i>Francesco Pinzin</i>	
CHAPTER 16	
Against control by implicit passive agents	279
<i>Lisa A. Reed</i>	

CHAPTER 17	
Romance evaluative <i>que/che/sã</i> sentences as inverted optatives	293
<i>Cristina Sánchez López</i>	
CHAPTER 18	
Resumed phrases (are always moved, even with in-island resumption)	309
<i>Dominique Sportiche</i>	
CHAPTER 19	
Timing properties of (Brazilian) Portuguese and (European) Spanish	325
<i>Irene Vogel, Angeliki Athanasopoulou and Natália Brambatti Guzzo</i>	
Language index	341
Subject index	343

Introduction LSRL 46 Stony Brook

The present volume emerges from the 46th annual Linguistic Symposium on Romance Languages (LSRL) that took place at Stony Brook University from March 31 to April 3, 2016. During this very successful event, linguists working on Romance languages had the opportunity to exchange ideas about different topics ranging from morphology, phonology, semantics, and syntax, to historical linguistics and sociolinguistics. Following the tradition of LSRL, here we provide a selection of papers presented at the event. Given the limitations imposed by the publisher of this volume, we had to leave out many excellent papers that were submitted. The criteria we followed in making a selection included quality as determined by a blind peer review process, as well as a balance of topics and languages represented. We sincerely thank the reviewers of the abstracts and the papers for their time and professionalism. Without their expertise this volume would not have been possible.

Various papers focused on information structure, left-right periphery, exclamatives, and *wh*-questions. Mayol and Villalba's contribution ("Bridging and dislocation in Catalan") discusses the relation of Right Dislocation and the notion of Bridging. Based on experimental results they show that speakers use Right Dislocation beyond contexts of Bridging. González and Reglero ("*Dime una cosa*: Are *wh*-in-situ questions different in Spanish? Evidence from intonation") show that echo-repetition and echo-surprise questions differ mostly in their intonational properties. These findings support a differential analysis between information-seeking questions and echo-questions on the one hand, and echo-repetition and echo-surprise questions on the other. In her study of "Morphological doublets in Brazilian Portuguese *wh*-constructions," Kato studies the distribution of *wh*-words from a diachronic and synchronic point of view. She studies Brazilian Portuguese *wh*-questions, and discusses the intriguing problem of apparent syntactic doublets, or optionality, in this domain. Exclamatives are studied by Sánchez-López ("Romance evaluative *que/che/sã* sentences as inverted optatives"). Her paper proposes that Romance evaluative sentences are expressive sentences with the overt complementizer that expresses the speaker's emotion about a proposition. She proposes a semantics based on bouletic alternatives and speakers' attitudes.

The morphology, semantics, and syntax of pronominals are also treated in various contributions. Charnavel studies the long-distance binding properties of the French pronoun *soi* in “Long distance binding of French reflexive *soi*: First-person oriented logophoricity”. She shows that *soi* is a standard anaphor that must be locally bound unless the relevant logophoric conditions are met. Maddox (“Licensing conditions on null generic subjects in Spanish”) studies Holmberg’s null generic subject generalization according to which only partial null subject languages allow null generic third-singular subjects, while consistent null subject languages do not. Maddox rejects Holmberg’s generalization based on the study of generic *uno* in Spanish and presents an alternative of the licensing conditions on null generics. The study of the syntax of resumptive pronouns is the subject of Sportiche’s contribution (“Resumed phrases (are always moved, even with in-island resumption)”). He concludes that resumption of a phrase by a pronominal element involves movement of the phrase in and outside island contexts. He proposes that movement takes place in two steps, the first being one of the Left Dislocation options, which feeds *wh*-movement or Left Dislocation. Reed (“Against Control by Implicit Passive Agents”) criticizes the idea that the agent of a passive verb is syntactically projected as a weak implicit argument. The article gives empirical arguments in support of an explanation in which the implicit agent of a passive verb is determined post-syntactically. Pescarini (“Stressed enclitics are not weak pronouns: A plea for allomorphy”) compares a morpho-phonological approach and a syntactic approach to stress shift and enclisis/proclisis asymmetries. The former approach assumes the same pronouns (allomorphs) are found in stress-shifting and non-stress-shifting contexts, while the latter assumes two series of pronouns belonging to different morphological classes. Pineda provides a nano-syntactic analysis of verbs of inherently directed motion with the presence of clitic *se* in “Causativization of verbs of directed motion in Romance languages”. Clitics are also the subject of Kayne’s contribution: “Clitic doubling, person and agreement in French hyper-complex inversion”. He analyzes hyper-complex inversion in which the object clitic agrees with the subject in some exceptional cases, providing an analysis in terms of clitic doubling.

Two articles examine negation. Déprez and Yeaton (“French negative concord and discord: An experimental investigation of contextual and prosodic disambiguation”) study double negation readings in negative concord languages by looking at context and prosody with new experimental data. Cépeda (“Expletive negation is not expletive: Evidence from aspect in Spanish”) looks at the semantics and syntax of negation in *until* clauses in Spanish and challenges the idea that Spanish punctual *hasta* clauses (‘until’) have identical meaning when containing (or not) an expletive negation. According to her proposal, negation in *hasta* clauses does in fact play a role in the meaning calculation.

Turning to semantics, Giorgi investigates the properties of counter-expectational surprise yes-no questions in Italian in her article “*Ma non era rosso?* (But wasn’t it red?): On counter-expectational questions in Italian”. She proposes that *ma* is a discourse head in a cartographic approach. Panaitescu (“Dependent numerals and dependent existentials in Romanian”) discusses Romanian dependent indefinites introduced by the distributive marker *câte*, and proposes a semantic analysis of their distribution.

In the area of morphology, Pinzin presents a historical study of Latin deponent verbs: “Latin denominal deponents: A syntactic analysis”. He shows that the use of deponents in Late Latin can be syntactically analyzed, contrary to purely morphological approaches to this phenomenon. Di Caro and Giusti study the inflected construction in Sicilian dialects (“Dimensions of variation: The inflected construction in the dialect of Delia (Caltanissetta)”), and provide an alternative syntactic analysis using ideas by Cardinaletti and Giusti on how functional projections interact in this dialect.

Vogel, Athanasopoulou, and Brambatti Guzzo test the hypothesis that syllable-timed languages (such as European Spanish) do not alter duration to express prominence, while non-syllable-timed languages (such as Brazilian Portuguese) do. Their findings, reported in their contribution “Timing properties of (Brazilian) Portuguese and (European) Spanish,” support the hypothesis and contribute to our understanding of rhythmic typology and the function of duration in a language’s prosodic system.

Guardiano, Michelioudakis, Cordoni, Irimia, Radkevich and Sitaridou (“Parametric comparison and dialect variation: Insights from Southern Italy”) apply micro-parameters in the nominal domain to show that the distribution of patterns agrees with the dialectal structure of Italo-Romance.

We would like to express our deepest gratitude to the following organizations for providing the financial support necessary to make this conference a success: the State University of New York (SUNY) for a “Conversations in the Disciplines” grant, the Stony Brook University Office of the Provost for a FAHSS (Faculty in the Arts, Humanities and Social Sciences) grant, as well as the Stony Brook University Center for Italian Studies, College of Arts and Sciences, Graduate Student Organization, and Department of Linguistics. The LSRL would not have been possible without the generous support and tireless contribution of time by the staff, graduate students, undergraduate students, and faculty in the Linguistics Department. Finally, special thanks to Olivia Mignone and Paola Cépéda for their assistance in the preparation of this volume.

Expletive negation is not expletive

Evidence from aspect in Spanish

Paola Céspedes

Stony Brook University / Pontificia Universidad Católica del Perú

Spanish punctual *hasta*-clauses (‘until’) have been described as identical in meaning when containing or not an expletive negation, that is, a negative marker that apparently does not contribute to the sentence meaning. I show that the negation in the *hasta*-clause does in fact play a role in the meaning calculation. My proposal is that the semantic contribution of the so-called “expletive” negation is to negate that the eventuality in the main clause holds during the interval denoted in the *hasta*-clause; as a result, such eventuality is expected to hold *after* the eventuality in the *hasta*-clause is finished. The novelty of my analysis is the comparison of the aspect (specifically, the durativity component) of the eventualities expressed in the *hasta*-clause.

Keywords: negation, expletive negation, until, accomplishment, achievement, temporal adverbial, Spanish, semantics

1. Introduction

In the literature, the label *expletive negation* (EN) is frequently used to refer to an apparently null negation that does not modify the truth value of the proposition it is included in because it does not contribute to the meaning of the sentence when appearing in certain specific environments. According to these assumptions, Spanish sentences (1) and (2) are treated as identical in meaning, even though the latter contains a negative marker in the *hasta*-clause, ‘until’ (Espinal 2000; Bosque 1980) (the *until*-clause appears between square brackets ‘[...]’):

- (1) *Olga no se irá [hasta que Daniela dé el discurso inaugural]*
Olga not CL will.go until that Daniela gives.SUBJ the speech opening
‘Olga won’t leave until Daniela gives the opening speech.’

- (2) *Olga no se irá [hasta que Daniela no dé el discurso*
 Olga not CL will.go until that Daniela NEG gives.SUBJ the speech
inaugural]
 opening
 ‘Olga won’t leave until Daniela gives the opening speech.’
 (Lit.: ‘Olga won’t leave until Daniela doesn’t give the opening speech.’)

The apparently vacuous nature of the negative marker *no* ‘not’ in the *hasta*-clause in (2) and other similar negative expressions in different languages has been described as a case of EN. According to this approach (which I will call the *expletive approach*), the negative marker is vacuous due to a process of logical absorption (Espinal 1992), or simply in virtue of having a non-negative meaning (Espinal 2000). The expletive approach, however, has some problems. If the negation did not contribute to the meaning of the sentence, it would be represented in logical form as the identity function (van der Wouden 1994). As Abels (2005) points out, it is implausible that the realization of the identity function and negation, two vastly different logical operators, should map onto the same morpheme.

Against the expletive approach, my goal in this paper is to show that the negative marker in the *hasta*-clause does contribute to the meaning calculation and, therefore, EN is not expletive at all. I claim that sentences (1) and (2) are semantically different with respect to the location of the main eventuality in time. The novelty of my proposal is the comparison of the aspect (specifically, the durativity component) of the eventualities expressed in the *hasta*-clause. In doing so, I find support for claiming that the truth conditions for a pair of sentences with and without EN are similar but not identical.

In Section 2, I discuss the different values of *until* and offer a semantics for Spanish punctual *hasta*-clauses. In Section 3, I provide support for the claim that EN does in fact contribute to the sentence meaning. I show that the expletive reading is just an illusion generated by the lack of durativity of an eventuality expressed in the *hasta*-clause. The paper concludes that, by conveying a negative meaning, the presence of EN changes the time of actualization of the eventuality in the main clause, and that EN is not expletive or logically vacuous after all.

2. The semantics of *hasta*-clauses

In this section, I separate atelic *hasta* from punctual *hasta*, and offer the pieces to calculate the meaning of a sentence containing punctual *hasta*. I apply this to the semantic analysis of punctual *hasta*-clauses.

2.1 The semantic components of punctual *hasta*

Karttunen (1974) proposes that there are two different values of *until*, which are applicable to the temporal readings of Spanish *hasta*. One value selects a salient time interval in which an eventuality is transpiring; I will call this “atelic *hasta*”. The other marks the time in which an eventuality actually takes place; I will call this “punctual *hasta*”.¹

An eventuality is compatible with atelic *until* only if it does not have well-defined termination points (Hitzeman 1991). In other words, atelic *until* combines with atelic predicates only, such as *correr* ‘to run’ (3a); it cannot combine with telic predicates, no matter whether they are durative (3b) or punctual (3c).

- (3) a. *Olga correrá hasta la medianoche.*
 Olga will.run until the midnight
 ‘Olga will run until midnight.’
 b. **Olga correrá dos millas hasta la medianoche.*
 Olga will.run two miles until the midnight
 c. **Olga llegará hasta la medianoche.*
 Olga will.arrive until the midnight

In contrast, punctual *until* locates punctual eventualities in time (Karttunen 1974, Giannakidou 2002, Condoravdi 2002). Therefore, it is compatible with negated telic predicates, as shown in (4):

- (4) *Olga no llegará hasta la medianoche.*
 Olga not will.arrive until the midnight
 ‘Olga won’t arrive until midnight.’

In the rest of this paper, I will focus on punctual *hasta* only, either followed by a temporal DP (e.g. *hasta la medianoche* ‘until midnight’) or by a clause (e.g. *hasta que Daniela dé el discurso inaugural* ‘until Daniela gives the opening speech’). For expository purposes, I will hereafter refer to punctual *hasta* as just *hasta*, except when a difference between atelic and punctual *hasta* is necessary.

Building on the proposals by Karttunen (1974), Giannakidou (2002) and Condoravdi (2002), I propose that the semantics of a sentence containing *hasta* is the result of the interaction of three components: an assertion, a factivity inference, and a scalar implicature.

By uttering a sentence containing *hasta*, the speaker commits herself to *assert* that the eventuality expressed in the main clause does not hold before the time

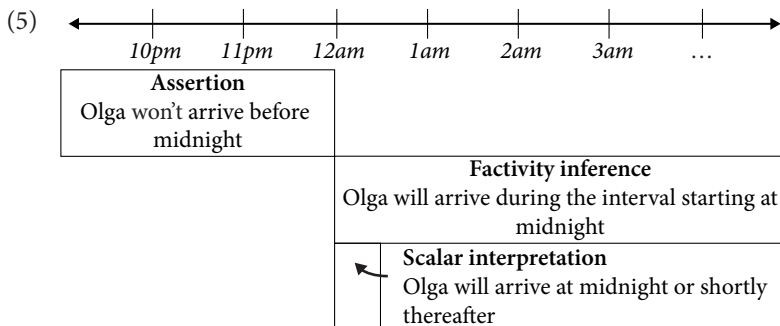
1. The literature has entertained different possibilities to explain the different values of *until* and its counterparts in different languages (e.g. Mittwoch 1977, de Swart 1996, a.o. vs. Condoravdi 2002, Giannakidou 2002, a.o.). This discussion, however, is out of the scope of this paper.

expressed by *hasta* has been reached. Accordingly, sentence (4) asserts that Olga will not arrive before midnight.

In addition, the eventuality in the main clause *has* to hold once the time explicitly expressed by *hasta* has been reached. I call this property the *factivity inference*, which cannot be canceled (Bassa Vanrell 2015). Thus, sentence (4) triggers the inference that Olga will definitely arrive and, because of this, it is inconsistent with Olga not arriving at all. Besides, the time expressed by *hasta* is the starting time of the interval in which the eventuality in the main clause is expected to hold. In this sense, not only does sentence (4) asserts that Olga will not arrive before midnight, but it also triggers the factivity inference that Olga will definitely arrive and midnight is the time at the very beginning of an interval in which Olga's arriving will occur.

An even stronger interpretation is that the time explicitly mentioned by *hasta* is the exact time in which the eventuality in the main clause actually holds. This property is the *scalar interpretation*. It is scalar because, of all the possible times included in the interval that starts at the time explicitly expressed by *hasta*, only this one is salient and there is no reason to believe that a different time is more relevant. In the case of sentence (4), the scalar interpretation is that Olga will arrive *at* midnight (or shortly thereafter).

A graphic representation of the three components of the meaning calculation of sentence (4) appears in (5):



2.2 Interpreting *hasta*-clauses

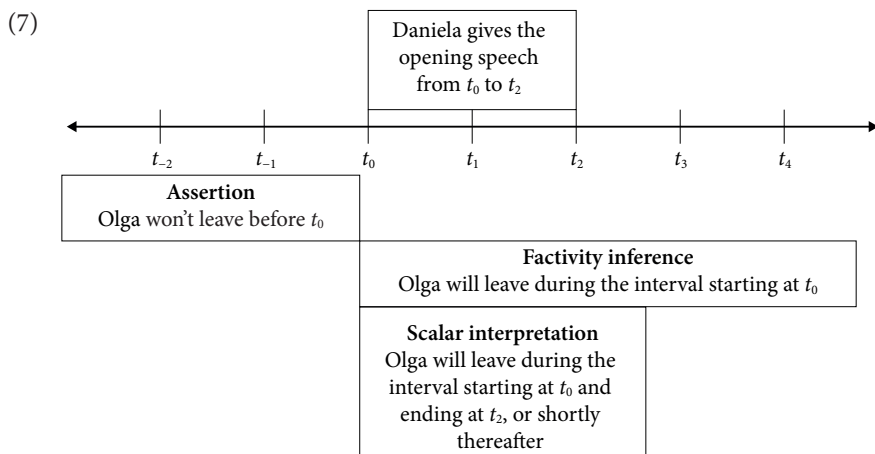
In *hasta*-phrases, the temporal indication is expressed by a temporal DP (e.g. *hasta la medianoche* 'until midnight'). In contrast, in *hasta*-clauses, the temporal indication seems to be related to an eventuality (e.g. *hasta que Daniela d e el discurso inaugural* 'until Daniela gives the opening speech'). More precisely, the temporal indication in *hasta*-clauses is the temporal trace of the eventuality expressed in it, that is, the time portion that such eventuality takes to occur.

Hasta-clauses frequently have a counterpart containing EN. As Bosque (1980) points out, they are probably the most common Spanish construction in which EN is allowed. For this reason, I will concentrate on the analysis of *hasta*-clauses.

Let us start by analyzing sentence (1), which contains an *hasta*-clause and is repeated here as (6):

- (6) *Olga no se irá* [hasta que Daniela dé el discurso inaugural]
 Olga not CL will.go until that Daniela gives.SUBJ the speech opening
 ‘Olga won’t leave until Daniela gives the opening speech.’

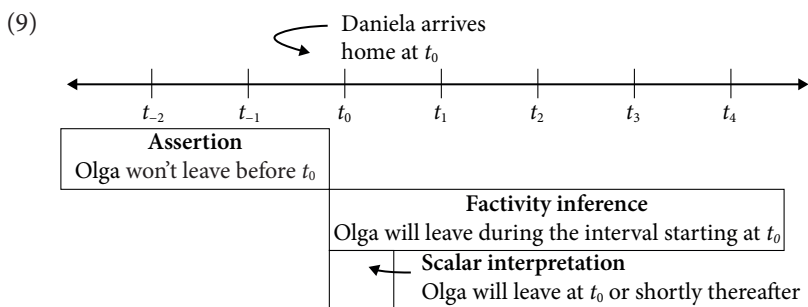
Note that, in (6), the temporal indication expressed in the *hasta*-clause corresponds to the time portion in which Daniela gives the opening speech. This eventuality is telic and extends over more than one time portion on a timeline (i.e. it is durative). Vendler (1957) calls eventualities with these properties *accomplishments*. Let us assume that, as an accomplishment, Daniela gives the opening speech from t_0 to t_2 . If so, then the basic assertion of sentence (6) is that Olga will not leave before t_0 (i.e. before Daniela starts giving the opening speech), the factivity inference is that Olga will leave during the interval starting at t_0 (i.e. once Daniela has started giving the opening speech), and the scalar interpretation is that Olga will leave during the relevant interval, the one starting at t_0 and ending at t_2 (i.e. while Daniela is giving the opening speech), or shortly thereafter. A graphic representation of this informal description is shown in (7):



Along with accomplishments, *hasta*-clauses can also contain *achievements*. Vendler (1957) characterizes them as telic, non-durative eventualities (i.e. they do not extend over more than one time portion on a timeline). An example appears in (8), where the eventuality of arriving home is an achievement.

- (8) *Olga no se ir  [hasta que Daniela llegue a casa]*
 Olga not CL will.go until that Daniela arrives.SUBJ to house
 ‘Olga won’t leave until Daniela arrives home.’

Assuming that Daniela arrives home at t_0 (as an achievement, it does not extend beyond one time portion), then sentence (8) asserts that Olga will not leave before t_0 (i.e. before Daniela arrives home), the factivity inference is that Olga will leave during the interval starting at t_0 (i.e. once Daniela has arrived home), and the scalar interpretation is that Olga will leave at t_0 (i.e. when Daniela arrives home), or shortly thereafter. A graphic representation of this informal description is shown in (9):



A formalization for the interpretation of sentences with *hasta*-clauses appears in (10). P and Q are the linguistic expressions of the eventualities in the main clause and the *hasta*-clause, respectively; t is a variable for a time interval, and t' for a contextually-determined interval; $<$ expresses precedence, and $\tau(e)$ is the temporal trace of the eventuality e .

- (10) *Semantics for sentences with hasta-clauses*
- a. Assertion:
 $\exists t \exists e [Q(t) \wedge P(e) \wedge \neg \exists t' [t' < t \wedge \tau(e)(t')]]$
 - b. Factivity inference:
 $\exists t \exists e [Q(t) \wedge P(e) \wedge \exists t' [t \leq t' \wedge \tau(e)(t')]]$
 - c. Scalar interpretation:
 $\exists t \exists e [Q(t) \wedge P(e) \wedge \exists t' [t \approx t' \wedge \tau(e)(t')]]$

The assertion in (10a) states that there exists a time interval t expressed in the *hasta*-clause and an eventuality e expressed in the main clause, and there does not exist any contextually-determined interval t' preceding t and related to the temporal trace τ of the eventuality e . The factivity inference in (10b) adds that there exists a contextually-determined interval t' that t is equal to or precedes, and related to the temporal trace τ of the eventuality e . Finally, the scalar interpretation in (10c) specifies that there exists a contextually-determined interval t'

approximately equal to t (representing the ‘shortly thereafter’ part of the informal description), and related to the temporal trace τ of the eventuality e .

As shown in (7) and (9), the left edge of the intervals in the factivity inferences and scalar interpretations exactly coincide with the left edge of the temporal indication in the *hasta*-clause. We will get back to this in the next section.

3. The negation in *hasta*-clauses

In this section, I bring up the problems that arise in some of the previous descriptions of EN in *hasta*-clauses. I also show that a sentence with EN in the *hasta*-clause receives a different interpretation from a sentence without it. In accordance to this, I offer a semantics for EN.

3.1 The distribution of EN in *hasta*-clauses

Bosque (1980) describes EN in *hasta*-clauses in a sentence such as (11), which repeats (2), as redundant. According to this author, the occurrence of *no* ‘not’ in the *hasta*-clause is conditioned either by the presence of a negative marker in the matrix clause or by a semantic overlapping with *mientras* ‘while’. Similarly, the Real Academia de la Lengua Española (2005) describes EN as a reinforcement of the matrix sentence negation, a reinforcement the sentence does not really need.

- (11) *Olga no se irá [hasta que Daniela no dé el discurso*
Olga not CL will.go until that Daniela NEG gives.SUBJ the speech
inaugural]
 opening
 ‘Olga won’t leave until Daniela gives the opening speech.’
 (Lit.: ‘Olga won’t leave until Daniela doesn’t give the opening speech.’)

These descriptions are misleading in at least two respects. First, Bosque’s (1980) statement that there may be a semantic overlapping between *hasta* ‘until’ and *mientras* ‘while’ is not accurate. These lexical items have different semantics. As shown by the grammaticality judgments of (12) and (13), *mientras*, requires an imperfective predicate (just like English *while*). Besides, in a sentence containing *mientras*, the eventuality in the main clause is not necessarily expected to hold, which contrasts with the factivity inference triggered by *until* (see Section 2.1.). Lexical items in other languages related to *until*, however, seem to have the semantics of *while*, as Russian *poka* (Abels 2005).

- (12) **Olga no habló* [*mientras que el bebé no nació*]
 Olga not spoke while that the baby not was.born
 Intended: ‘Olga didn’t say a word while the baby wasn’t born.’
- (13) *Olga no habló* [*hasta que el bebé no nació*]
 Olga not spoke until that the baby NEG was.born
 ‘Olga didn’t say a word until the baby was born.’
 (Lit.: ‘Olga didn’t say a word until the baby wasn’t born.’)

Second, both Bosque (1980) and the Real Academia de la Lengua Española (2005) coincide in that EN may appear in *hasta*-clauses when there is a negative marker in the matrix. However, this is not the only context for EN to occur in these constructions. When the matrix clause has continuative sense and there is no negation in it, an atelic *hasta*-clause may also contain EN. This is shown in the following examples:

- (14) *Seguiré gritando* [*hasta que no me des mi dinero*]
 will.continue yelling until that NEG me give.SUBJ my money
 ‘I’ll keep on yelling until you give me my money.’
 (Lit.: ‘I’ll keep on yelling until you don’t give me my money.’)
- (15) *Es inocente* [*hasta que no se demuestre lo contrario*]
 is innocent until that NEG CL proves.SUBJ the contrary
 ‘She is innocent until proven guilty.’
 (Lit.: ‘She is innocent until they don’t prove the contrary.’)

The grammaticality of the sentences (14) and (15) suggests that, contra Bosque (1980) and the Real Academia de la Lengua Española (2005), the idea of redundancy or reinforcement is not the explanation for the presence of *no* in *hasta*-clauses.

This is further supported by the fact that *hasta*-clauses may contain a so-called “ordinary” or “real” negation (i.e. a negative marker that does in fact change the truth value of the proposition). Thus sentence (16) is ambiguous: the negation in the *hasta*-clause can be interpreted either as “ordinary” (16a) or “expletive” (16b):

- (16) *No volveremos* [*hasta que no haya agua en nuestra casa*]
 not will.return until that NEG has.SUBJ water in our house
- a. “Ordinary” negation reading:
 ‘We won’t come back until there’s no water in our house.’
 - b. “Expletive” negation reading:
 ‘We won’t come back until there’s water in our house.’

For reading (16a), one can imagine a situation in which some pipes broke and the house got flooded. In this context, we will come back only when the water has

been drained off. In contrast, for reading (16b), the situation can be one in which there is no water in the house, so we will come back only when water is back.

Since reading (16a) is available, it is at least suspicious to assume that the negation in the *hasta*-clause is simply a reinforcement strategy: it cannot be the case that it is “ordinary” negation in one case but it completely loses its meaning in the other just to serve as a reinforcement marker.

An additional piece of data is offered by sentences such as (17). The negative marker as presented in this example can license a negative polarity item (NPI) such as *ni un centavo*, ‘(not) even a red cent’:

- (17) *Olga no se irá [hasta que Daniela no tenga ni un centavo]*
 Olga not CL will.go until that Daniela NEG has.SUBJ even a cent
 ‘Olga won’t leave until Daniela doesn’t have a red cent.’

The sentence in (17) is not ambiguous. It means that Olga will leave only when Daniela runs out of money. It is simply not possible to have a second reading in which the negation is a reinforcement marker. An “expletive” reading is just out in this context. An account for EN should explain why the “expletive” reading is not available in the presence of a concomitant NPI.

3.2 The semantic contribution of EN

As I said in Section 1, sentences (1) and (2), repeated here as (18) and (19), respectively, have been claimed to be identical in meaning, even though the latter contains EN in the *hasta*-clause. However, these sentences have different truth conditions as one of them can be true while the other is false.

- (18) *Olga no se irá [hasta que Daniela dé el discurso inaugural]*
 Olga not CL will.go until that Daniela gives.SUBJ the speech opening
 ‘Olga won’t leave until Daniela gives the opening speech.’

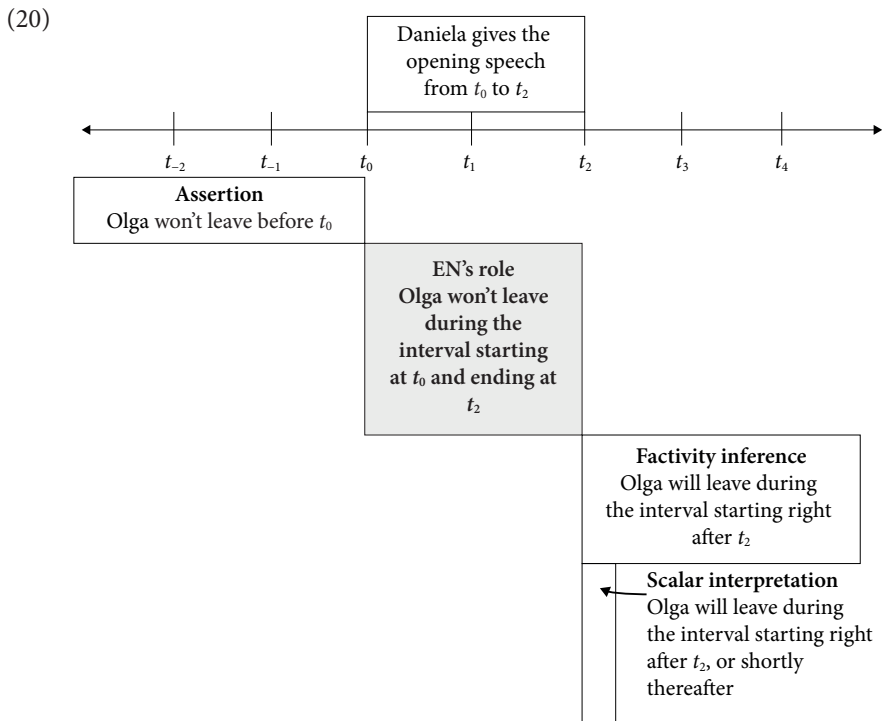
- (19) *Olga no se irá [hasta que Daniela no dé el discurso inaugural]*
 Olga not CL will.go until that Daniela NEG gives.SUBJ the speech opening
 ‘Olga won’t leave until Daniela gives the opening speech.’
 (Lit.: ‘Olga won’t leave until Daniela doesn’t give the opening speech.’)

Let us imagine the following situation. Olga wants to leave once Daniela starts giving the opening speech or while she is saying the first words at the latest. In this situation, (18) is true, but (19) is false. Sentence (19) expresses that Olga will make sure that Daniela delivers the *whole* opening speech and she will leave only after

Daniela has finished (i.e. when Daniela is *no longer* giving the opening speech). Sentence (18), however, is not a guarantee for that.

The different truth conditions of sentences with and without EN suggest that EN is actually playing a negative role. It negates that the eventuality in the main clause will hold during the interval denoted in the *hasta*-clause, and restricts the factivity inference to the complement of said interval. In consequence, the eventuality in the main clause is expected to hold *after* the temporal indication in the *hasta*-clause is finished.

A graphic representation of the effect of EN in sentence (19) is shown in (20):



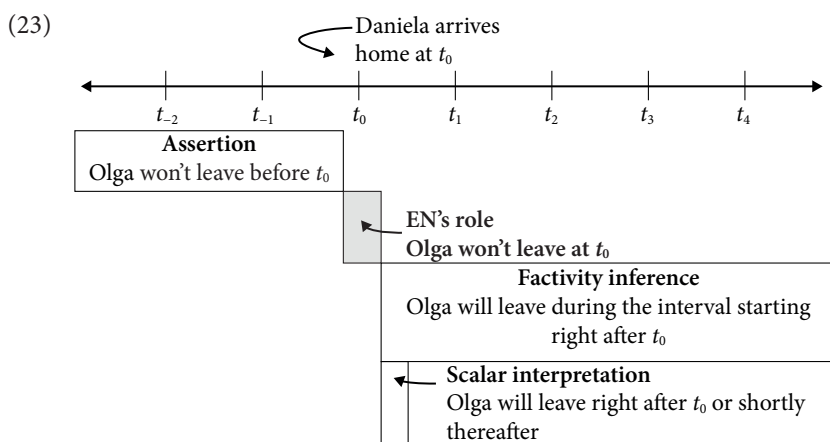
When comparing the graphic representations in (7) and (20), we can see that sentences without EN and sentences with EN express a different location of the main eventuality in time. In the former, Olga's leaving will hold at t_0 or shortly thereafter, whereas in the latter Olga's leaving will hold right after t_2 or shortly after that. In other words, while in (7) the left edge of the intervals in the factivity inference, the scalar interpretation and the *hasta*-clause exactly coincide, this does not apply in (20), where the left edge of the intervals in the factivity inference and scalar interpretation coincide with the right edge of the interval in the *hasta*-clause (i.e. its end).

I have exemplified the role of EN with sentence (19). Its *hasta*-clause contains an accomplishment, which makes it clear to show that the eventuality in the main clause will hold when the interval denoted in the *hasta*-clause is no longer running. However, when *hasta*-clauses contain an achievement, the role of EN is not as evident. Let us consider the examples in (21), which repeats (8), and its counterpart with EN in (22):

(21) *Olga no se irá [hasta que Daniela llegue a casa]*
 Olga not CL will.go until that Daniela arrives.SUBJ to house
 ‘Olga won’t leave until Daniela arrives home.’

(22) *Olga no se irá [hasta que Daniela no llegue a casa]*
 Olga not CL will.go until that Daniela NEG arrives.SUBJ to house
 ‘Olga won’t leave until Daniela arrives home.’

An informal description for sentence (21) was offered in (9). According to it, Olga will leave when Daniela arrives home or shortly thereafter, which makes the left edge of the intervals in the factivity inference, the scalar interpretation and the *hasta*-clause exactly coincide. In contrast, a different result appears when analyzing (22), whose *hasta*-clause contains an achievement and EN. Following my proposal for the semantic contribution EN and assuming that Daniela arrives home at t_0 , the role of EN in sentence (22) is to negate that Olga will leave during t_0 (i.e. while Daniela is arriving home). A graphic representation of this role is shown in (23):



Now, when comparing the graphic representations in (9) and (23), we can see that, without EN, the sentence expresses that Olga will leave when Daniela arrives home or shortly thereafter, whereas with EN, the sentence expresses that Olga will leave immediately after Daniela arrives home or shortly after that. However, the intervals in the factivity inferences and the scalar interpretations have left limits so

close in (9) and (23) that both interpretations are very similar, as they may cover the same time portion if we take the ‘shortly thereafter’ part into consideration. As a result, speakers are inclined to think that the negative marker is vacuous and does not play any role in the meaning calculation of the sentence.

As I claimed in a previous work (C epeda 2017), this is just an *illusion of expletiveness*. The illusion of expletiveness is created when *hasta*-clauses contain an achievement. Because of the lack of durativity of an achievement, the role of EN is not as straightforward as it is when combined with accomplishments. However, either with accomplishments or achievements, EN is actually changing the time location of the eventuality in the main clause by changing the left edge of the intervals in the factivity inference and scalar interpretation.

Following the previous discussion, a formalization for the interpretation of sentences containing EN in the *hasta*-clause is proposed in (24). The symbols follow the legend outlined for (10); along with them, \prec_i means ‘immediately precedes’ and the contribution of EN is underlined in (24a):

(24) *Semantics for sentences with hasta-clauses + EN*

a. Assertion:

$$\exists t \exists e [Q(t) \wedge P(e) \wedge \neg \exists t' [t' \prec t \wedge \tau(e)(t')]] \wedge \underline{\neg [\tau(e)(t)]}$$

$$\Leftrightarrow \exists t \exists e [Q(t) \wedge P(e) \wedge \neg \exists t' [t' \preceq t \wedge \tau(e)(t')]]$$

b. Factivity inference:

$$\exists t \exists e [Q(t) \wedge P(e) \wedge \exists t' [t \prec t' \wedge \tau(e)(t')]]$$

c. Scalar interpretation:

$$\exists t \exists e [Q(t) \wedge P(e) \wedge \exists t' [t \prec_i t' \wedge \tau(e)(t')]]$$

The assertion on the first line in (24a) contains the one in (10a) and it further states that it is not the case that the temporal trace τ of the eventuality e expressed in the main clause corresponds to t (which is the interval expressed in the *hasta*-clause). The second line in (24a) is a simplification of the first line to express that there does not exist any contextually-determined interval t' equal to or preceding t related to temporal trace τ of the eventuality e . The factivity inference in (24b) states that there exists a contextually-determined interval t' preceded by t , and related to the temporal trace τ of the eventuality e . Finally, the scalar interpretation in (24c) specifies that there exists a contextually-determined interval t' immediately preceded by t , and related to the temporal trace τ of the eventuality e .

As shown in the first line of the assertion in (24a), EN has very narrow scope: it only scopes over the temporal trace function; it does not have scope over any part of the linguistic content of the eventuality. EN’s narrow scope explains the unavailability of the ‘expletive’ reading when a strong NPI needs to be licensed by negation, as shown in example (17). EN cannot license NPIs, which suggests that its scope is very restricted in the clause. Note that this inability does not support

the claims about EN being a reinforcement strategy, contrary to Bosque's (1980) and the Real Academia de la Lengua Española's (2005) positions.

To sum up, I have proposed that the role of EN in Spanish *hasta*-clauses is indeed negative, as it negates that the eventuality in the main clause will hold while the temporal indication in the *hasta*-clause is transpiring. The assertion of a sentence with EN in the *hasta*-clause entails the assertion of one without EN in it, but not inversely. Consequently, EN changes the size of the time interval in which an eventuality is expected to hold. I want to claim that EN does also play a role in the delimitation of the interval expressed in other temporal adverbial clauses in languages different from Spanish, such as German *bevor*-clauses 'before' (Krifka 2010) or Korean *-ci*-clauses 'since' (Yoon 2011), where EN also occurs.

In general, it seems that, by conveying a negative meaning, EN has the effect of changing a domain of quantification. This effect may be found in contexts different from the temporal domain in which EN also appears, such as comparatives (Donati 2000), *nearly*-clauses (Schwenter and Pons Bordería 2005), free relatives (Eilam 2008), subordinated clauses of predicates of prohibition or fear (van der Wouden 1994), exclamatives (Portner and Zanuttini 2000), and so on. In the case of *hasta*-clauses, as I have shown, the role of EN becomes evident when contrasting the aspect, specifically the durativity component, of the eventuality in the *hasta*-clause. However, in other cases, the role of EN may also be manifested by other means.

4. Conclusions

I have shown that, against the expletive approach, Spanish sentences containing *hasta*-clauses with and without EN are semantically different. EN is not semantically null, but does actually have a negative meaning. It changes the time of actualization of the eventuality in the main clause by negating that it will occur during the temporal interval expressed in the *hasta*-clause. Therefore, the truth conditions of the sentences are not the same, which supports the claim that EN is not expletive after all.

Acknowledgements

✉ paola.cepada@stonybrook.edu, pcepada@pucp.edu.pe

I want to express my gratitude to Richard Larson, Jiwon Yun, Viviane Déprez, Francisco Ordóñez, Gary Mar, Thomas Graf, Michael Becker, Marie Huffman, Robert Pasternak, Gabriel Martínez-Vera, Chong Zhang, So Young Lee, Yaobin Liu, Lei Liu, Hongchen Wu, Nazila Shafiei,

Vahideh Rasekhi, the audiences at TLS16, PLC40, LSRL46, CLS52, Stony Brook University and Pontificia Universidad Católica del Perú, and the anonymous reviewers for helpful discussion and comments. A previous version of this article appeared as Céspedes (2017).

References

- Abels, Klaus. 2005. "‘Expletive Negation’ in Russian: A Conspiracy Theory." *Journal of Slavic Linguistics* 13: 5–74.
- Bassa Vanrell, María del Mar. 2015. "One or two UNTILs? The case of single-UNTIL languages." Presentation at MIT LingLunch.
- Bosque, Ignacio. 1980. *Sobre la negación*. Madrid: Cátedra.
- Céspedes, Paola. 2017. "Against expletive negation: The case of Spanish *hasta*-clauses." In *Proceedings of the 52nd Annual Meeting of the Chicago Linguistics Society (2016)*, ed. by Jessica Kantarovich, Tran Truong, and Orest Xherija, 101–115. Chicago: CLS.
- Condoravdi, Cleo. 2002. "Punctual *until* as a scalar NPI." In *The nature of the word. Studies in honor of Paul Kiparsky*, ed. by Kristin Hanson, and Sharon Inkelas, 63–654. Cambridge, MA: MIT Press.
- De Swart, Henriëtte. 1996. "Meaning and use of *not... until*." *Journal of Semantics* 13: 221–263. <https://doi.org/10.1093/jos/13.3.221>
- Donati, Caterina. 2000. "A note on negation in comparison." *Quaderni del Dipartimento di Linguistica-Università di Firenze* 10: 55–68
- Eilam, Aviad. 2008. "The Crosslinguistic Realization of *-ever*. Evidence from Modern Hebrew." Ms., University of Pennsylvania.
- Espinal, María Teresa. 1992. "Expletive negation and logical absorption." *The Linguistic Review* 9: 333–358. <https://doi.org/10.1515/tlir.1992.9.4.333>
- Espinal, María Teresa. 2000. "Expletive Negation, Negative Concord and Feature Checking." *Catalan Working Papers in Linguistics* 8: 47–69.
- Giannakidou, Anastasia. 2002. "UNTIL, Aspect, and Negation: A Novel Argument for Two *Untils*." In *Proceedings of SALT XII*, ed. by Brendan Jackson, 84–103. Ithaca, NY: Cornell University.
- Hitzeman, Janet. 1991. "Aspect and Adverbials." In *Proceedings of SALT I*, ed. by Steven K. Moore, and Adam Zachary Wyner, 107–126. Ithaca, NY: Cornell University.
- Karttunen, Laurie. 1974. "Until." In *Papers from the Tenth Regional Meeting, Chicago Linguistic Society*, ed. by Michael W. La Galy, Robert A. Fox, and Anthony Bruck, 284–297. Chicago: CLS.
- Krifka, Manfred. 2010. "How to interpret ‘expletive’ negation under *bevor* in German." In *Language and logos. Studies in theoretical and computational linguistics*, ed. by Thomas Hanneforth, and Gisbert Fanselow, 214–236. Berlin: Akademie Verlag. <https://doi.org/10.1524/9783050062365.214>
- Mittwoch, Anita. 1977. "Negative sentences with *until*." In *Papers from the 13th Regional Meeting, Chicago Linguistic Society*, ed. by Woodford A. Beach, 410–417. Chicago: CLS.
- Portner, Paul, and Raffaella Zanuttini. 2000. "The force of negation in *wh*-exclamatives and interrogatives." In *Negation and polarity: Syntactic and semantic perspectives*, ed. by Lawrence R. Horn, and Yasuhiko Kato, 201–39. New York: Oxford University Press.
- Real Academia de la Lengua Española. 2005. *Diccionario Panhispánico de Dudas*. Madrid: RAE.

- Schwenter, Scott, and Salvador Pons Bordería. 2005. “*Por poco (no)*: Explicación sincrónica y diacrónica de sus componentes de significado.” *Lingüística Española Actual* XXVIII: 131–158.
- Vendler, Zeno. 1957. “Verbs and times.” *Philosophical Review* 56: 143–160.
<https://doi.org/10.2307/2182371>
- van der Wouden, Ton. 1994. “Polarity and ‘illogical negation.’” In *Dynamics, polarity, and quantification*, ed. by Makoto Kanazawa, and Christopher J. Piñón, 17–45. Stanford: CSLI.
- Yoon, Suwon. 2011. “‘Not’ in the mood. The syntax, semantics and pragmatics of evaluative negation.” Ph.D. dissertation, University of Chicago.

Long-distance binding of French reflexive *soi*

First-person oriented logophoricity

Isabelle Charnavel
Harvard University

This paper aims at showing that long-distance binding can be reduced to logophoric exemption from Condition A based on the case study of French *soi*. *Soi* is usually treated as a long-distance anaphor similar to Icelandic *sig*. But once the relevant factors are disentangled, it turns out that *soi* is a standard anaphor that must be locally bound unless the relevant logophoric conditions are met, which exempt *soi* from binding requirements. First, *soi* is a first-person oriented generic, expressing a generalization based on the discourse participants' identification with the antecedent. Second, this determines its logophoric conditions of exemption (speaker-orientedness): *soi* is exempt from Condition A when its antecedent includes the speaker and is the perspective center of its clause.

Keywords: anaphor, Condition A, long-distance binding, logophoricity, speaker orientation, genericity, French

1. Introduction

Anaphors like English *himself* are standardly assumed to be subject to locality constraints with respect to their antecedents (Condition A of the Binding Theory; Chomsky 1986, a.o.). In various language families, it has nevertheless been observed that some instances of anaphors do not appear to obey Condition A, as illustrated in (1)–(2): under any version of locality, English *himself* and Mandarin *ziji* do not have local antecedents.

- (1) *Max_i boasted that the queen invited Lucie and himself_i for a drink.*
(Reinhart and Reuland 1993: 670)
- (2) *Zhangsan_i shuo Lisi piping-le ziji_i*
'Zhangsan_i said Lisi criticized him_i.'
(Huang and Liu 2001: 13)

Two types of hypotheses have been made to account for such cases. On the one hand, it has been claimed that some anaphors (simplex anaphors like Icelandic *sig*, see Reuland 2006, a.o., or Mandarin *ziji*, see Cole et al. 2006, a.o.) can be long distance bound, i.e. the binding domains of anaphors should be parameterized (Manzini and Wexler 1987, a.o.) and the domain size correlates with the morphological complexity of anaphors. On the other hand, it has been proposed that some anaphors (complex anaphors like English *himself*, see Pollard and Sag 1992, a.o., but also simplex anaphors like Icelandic *sig*, see Maling 1984, a.o., or Mandarin *ziji*, see Huang and Liu 2001, a.o.) can be exempt from Condition A under logophoric conditions, i.e. when occurring in clauses expressing the perspective of their antecedent (Sells 1987, a.o.).

This twofold treatment of anaphors disobeying Condition A seems problematic. Empirically, it is questioned by the behavior of simplex anaphors like Icelandic *sig* or Mandarin *ziji* that do not have to be bound under logophoric conditions and often (perhaps always) need to be logophoric to be long distance bound (Maling 1984, Huang and Liu 2001, a.o.). Theoretically, it raises an issue of parsimony, as anaphors end up being subject to various possible constraints (binding within a local domain, an intermediate domain or an unbounded domain, or absence of binding), crosslinguistically or even within the same language.

For these reasons, it seems more promising to adopt a unified treatment of anaphors disobeying Condition A in terms of logophoricity (cf. Charnavel and Sportiche 2016a): anaphors are either bound within a uniform local domain or they are exempt from locality requirements under perspective-related conditions.

In this article, I will use French reflexive *soi* (\approx ‘oneself’) as a case study to support this hypothesis. *Soi* has been treated as a simplex reflexive falling into the category of long distance anaphors like Icelandic *sig* (Pica 1987, a.o.), which can be bound from outside their local domain. The goal of this paper is to show that *soi* is in fact a well-behaved anaphor that obeys Condition A unless the relevant logophoric conditions are met. What obscures its behavior, I will argue, is that *soi* is a first-person oriented generic (Section 2). Once this is taken into account, we can observe that *soi* is exempt from Condition A when the speaker is included in its antecedent and is the center of perspective (Section 3).

2. French *soi* as a first-person oriented generic

2.1 Background about *soi*

French *soi*, which derives from Latin accented *se* (and is thus related to French clitic *se*), is a non-clitic reflexive, unmarked for gender and number, and precluded

from nominative positions. According to Ronat (1982), *soi* obeys the standard Condition A and requires an indefinite, arbitrary antecedent as exemplified by the following sentences (from Ronat 1982: 193).

- (3) *Personne_i ne doit voter pour soi_i*
 ‘Nobody_i should vote for oneself_i.’¹
- (4) *On_i rêve souvent de soi_i*
 ‘One_i often dreams about oneself_i.’
- (5) *Il ne faut pas PRO_i douter de soi_i*
 ‘One_i should not have doubts about oneself_i.’
- (6) **Personne_i ne sait que [tu votes pour soi_i]*
 ‘*Nobody_i knows that [you vote for oneself_i].’
- (7) **Personne_i n’a laissé [ces gens voter pour soi_i]*
 ‘*Nobody_i let [these people vote for oneself_i].’

In (3)–(5), *soi* is bound by an arbitrary or indefinite antecedent within its smallest XP containing a subject. In (6)–(7), however, *soi* is bound from outside its minimal clause. According to Ronat (1982), this explains the difference in acceptability: *soi* obeys Condition A only in the first set of examples.

Based on the following examples, Pica (1984, 1987, 1991) argues against Ronat’s (1982) proposal and hypothesizes that just like Icelandic *sig* (see Reuland 2006, a.o.), *soi* can be long distance bound as long as its antecedent is within the smallest tensed clause containing it.

- (8) *On_i ne devrait jamais laisser [_{INF} les gens dire du mal de soi_i]*
 ‘One_i should never let people speak ill of oneself_i.’
- (9) *Personne_i ne souhaite jamais que [_{SUBJ} les gens disent du mal de soi_i]*
 ‘Noone_i ever wishes people speak ill of oneself_i.’
- (10) **Personne_i ne dit que [_{IND} les gens ont pensé à soi_i]*
 ‘*Noone_i is saying that people thought about oneself_i.’

Soi is acceptable when it is bound from outside its infinitive clause as in (8) or its subjunctive clause as in (9), but not when its antecedent sits outside its minimal indicative clause as in (10). According to Pica, this demonstrates that *soi* can be long distance bound: its binding domain is not the smallest XP with subject containing it, but its smallest tensed clause.

1. Throughout the article, *soi* is translated as ‘oneself’ because it is the closest English equivalent, even if *oneself* is not identical to *soi* in several respects. The judgments indicated in the translations are those of the French sentences, not those of the English counterparts.

One crucial issue with Ronat's and Pica's proposals, however, is that they do not take into account the genericity of *soi* (cf. Zribi-Hertz 1990: the antecedent of *soi* must be a universal quantifier). In particular, tense and adverbs favor a generic interpretation in (3)–(5) and (8)–(9), but not in (6)–(7) and (10). Once genericity is controlled for, it turns out that *soi* can be bound from outside its indicative clause as shown by (11) (Zribi-Hertz 1990: 121).

- (11) *On_i espère toujours que [_{IND} les autres ont dit du bien de soi_i]*
 'One_i always hopes that others spoke highly of oneself_i.'

In order to determine the binding properties of *soi*, it is thus crucial to determine the generic properties of *soi*, which can be done independently based on the uncontroversial cases of local binding.

2.2 Genericity of *soi*

The antecedent of *soi* exhibits several restrictions. First, it must be human as illustrated by (12) (Zribi-Hertz 1990: 120):²

- (12) *{Pas grand chose/rien_i} ne peut se recroqueviller sur soi_i*
 '{[Not much/nothing_i] can curl up on oneself_i.'

Second, the antecedent of *soi* must be non-specific, that is, unmarked for gender, arbitrary or indefinite.³ For instance, we have observed in (4)–(5) that impersonal *on* (≈ 'one') and arbitrary PRO can antecede *soi*. More specifically, impersonal *on* has several uses (Creissels 2008, a.o.), namely an existential/episodic use, a generic/gnomic use and a first-person plural use similar to *nous* 'we', and *soi* can only be anteceded by generic *on* as in (4), not by existential *on* as in (13) (adapted from Rezac and Jouitteau 2008) or first-person plural *on* as in (14).

- (13) *On_i m'a laissé un message sur le répondeur, mais on ne m'a rien dit (*sur soi_i)*
 'Someone_i left a message on my phone but didn't say anything (*about oneself_i).'

2. There is some dialectal variation: some French speakers accept non-human antecedents in some cases. The dialectal variation also affects other properties described here (some speakers may accept plural or existential antecedents). In this paper, I only consider the standard dialect (which is my dialect), where all this is impossible, and I leave for future research a comparison of the dialectal variations.

3. Historically, *soi* licensed specific human antecedents as in (i) (from Stendhal 1835). Some speakers (of non-standard dialects) still accept such uses.

- (i) *M. Leuwen_i fut content de soi_i*
 'M. Leuwen_i was happy about himself_i.'

- (14) *Nous aussi, on_i a parlé* (**de soi_i*)
 ‘We_i too spoke (*about oneself_i)’

Third, the antecedent of *soi* must be a universal quantifier (Zribi-Hertz 1990). That’s why *personne* ‘nobody’ can antecede *soi* in (3), but not *quelqu’un* ‘someone’ below (Zribi-Hertz 1990: 120).

- (15) **Quelqu’un_i aura sûrement pensé à soi_i*
 ‘*Someone_i will probably have thought of oneself_i.’

That’s also why only generic *on* is licensed with *soi*, as opposed to existential or first-person plural *on*.

Fourth, *soi* requires a singular antecedent as shown by the contrast in (16) (Kayne 2003, citing J.-Y. Pollock). Based on this, Kayne (2003) hypothesizes that *soi* contains two morphemes, namely *s-* (underspecified person) and *-oi* (singular) and forms a natural class with *t-oi* and *m-oi* (strong pronouns ‘you’ and ‘me’).

- (16) a. [*Tout linguiste*]_i parle de soi_i⁴
 ‘[Every linguist]_i talks about oneself_i.’
 b. * [*Tous les linguistes*]_i parlent de soi_i
 ‘*[All the linguists]_i talk about oneself_i.’

This also explains why *soi*, unlike *se*, does not exhibit reciprocal or collective readings, which would require plural antecedents:

- (17) #*Quand on_i parle mal de soi_p, ça peut dégénérer* [*reciprocal]
 ‘#When one_i speaks badly of oneself_p, things can go sour.’
 Intended reciprocal interpretation: when people speak badly of each other, things can go sour.
 (18) *En France, on_i est fier de soi_i* [*collective; distributive]
 ‘In France, one_i is proud of oneself_i.’
 Intended collective interpretation: ‘In France, people are proud of their nation.’

For all these reasons, I hypothesize that *soi* is a first-person generic in the sense of Moltmann (2006) based on English *one* (see also Malamud 2012): it expresses a generalization based on first-person experience. That is, the speaker uses *soi* when she identifies with – or empathizes with – the referent to convey a general statement: *soi* roughly means ‘people in general with whom I identify’.

4. While non-restricted genericity (e.g. *on* ‘one’) always licenses *soi*, restricted genericity (with domain restriction as in *tout linguiste* ‘every linguist’) gives rise to dialectal variations in judgments.

This does not mean that the speaker must be included in the generalization as shown by (19)–(20): both examples are acceptable, even if *ici* ‘here’ in (19) suggests that the speaker is included in the referent of *on*, but *là-bas* ‘over there’ in (20) suggests that she is not. But this means that the speaker must empathize with the participants, as demonstrated by the contrast between (20) involving *soi* and (21) involving the third person *eux* ‘them’, which implies that the speaker distances herself from the referents.

- (19) *Ici en France, on_i a confiance en soi_i*
 ‘Here in France, one_i has confidence in oneself.’
- (20) *Là-bas en Grèce, on_i garde confiance en soi_i*
 ‘Over there in Greece, one_i remains confident in oneself.’
- (21) *Là-bas en Grèce, [les gens]_i gardent confiance en eux_i*
 ‘Over there in Greece, people_i remain confident in themselves.’

Note that by using *soi*, the speaker also intends to include the addressee in the identification (cf. *tu/vous* ‘you’ for generics that are more specifically oriented towards the addressee), so that *soi* implies empathy of all discourse participants with the generic referent (*soi* ≈ ‘people in general with whom you and I identify’).

Moreover, *soi* can only be oriented towards the participants of the actual discourse, not those of a reported discourse. Specifically, unlike the generic variable hypothesized in the case of taste predicates (Pearson 2013), the empathic component of *soi* cannot shift from the speaker to the attitude holder.

- (22) #*Janis dit qu’[en Grèce, on_i garde confiance en soi_i]*
 ‘#Janis says that [in Greece, one_i remains confident in oneself].’
- (23) *John thinks that [this cake is tasty].*
 ≈ This cake is tasty to every individual with whom John identifies.

While the individual identifying with the participants in the generalization can shift to the attitude holder in (23), it cannot in (22): even assuming that the speaker does not know anything about Greece and that the attitude holder Janis is Greek, the generalization expressed in the bracketed embedded proposition cannot be based on Janis’ experience (although plausibility favors this interpretation), but has to be based on the speaker’s (in which case the sentence is perfectly acceptable).

The empathic component of *soi* is thus unshiftable, just like French first and second-person pronouns. Another similarity of *soi* with indexicals is that just like the first-person pronoun *moi* ‘me’, *soi* does not have to be read *de se* in attitude contexts, unless it is turned into an intensifier by the morpheme – *même* (lit. ‘same’), which introduces focus (cf. Zribi-Hertz 1990, a.o.).

- (24) *On_i rêve parfois qu'on est [un monstre]_k et qu'on_k lance une attaque contre {soi_{i/k} / soi_{*i/k}-même}*
 'One_i sometimes dreams that one is [a monster]_k and one_k launches an attack against {oneself_{i/k} / oneself_{*i/k}-same}.'
- (25) *Je_i rêve parfois que je suis [un monstre]_k et que je_k lance une attaque contre {moi_{i/k} / moi_{*i/k}-même}*
 'I_i sometimes dream that I am [a monster]_k and I_k launch an attack against {me_{i/k} / myself_{*i/k}}.'

Soi in (24), just like *moi* in (25), can refer to the *de re* dreamer, while *soi-même*, just like *moi-même*, must refer to the *de se* monster (as indicated by the different possibility of indices).

In sum, *soi* has several components: (i) an indexical component (arguably expressed by the morpheme *s-*) accounting for its humanness and its orientation towards discourse participants; (ii) a singular feature (arguably expressed by the morpheme *-oi*) explaining the impossibility of plurally marked antecedents and that of reciprocal and collective readings; (iii) a generic component, which can be represented as a binding requirement by a generic operator (cf. Chierchia 1998, a.o.). Importantly, the relation between components (i) and (iii), which is made possible by the underspecification in person, is complex: discourse participants are not necessarily included in the generic antecedent of *soi*, but they identify with it (cf. 'generic simulation' in Moltmann 2006). It is in that sense that *soi* is a first-person generic.

3. French *soi* as an anaphor exempt under logophoric conditions

3.1 Exemption of *soi*

Now that we have specified the generic properties of *soi*, we can go back to its binding properties. Recall that we have observed that *soi* can be locally bound (Ronat 1982, a.o.) and that it can also be bound from outside infinitive and subjunctive clauses (Pica 1984, 1987, 1991 vs. Ronat 1982), as well as indicative clauses when genericity is controlled for (Zribi-Hertz 1990 vs. Pica 1987). The following example (from Sportiche et al. 2013: 435) further illustrates this.

- (26) *Dans ces cas-là, on_i ne se rend pas compte que la télé parle de soi_i*
 'In such cases, one_i does not realize that the TV talks about oneself_i.'

Furthermore, *soi* does not even need to be bound as shown in (27), where the antecedent *on* does not c-command *soi*, and in (28)–(29), where *soi* does not have any antecedent in the sentence.⁵

(27) *Le mal qu'on_i inflige à autrui peut se retourner contre soi_i.*
 'The harm one_i does to others can turn against oneself_i.'

(28) *Le repli sur soi peut conduire à la dépression*
 'Withdrawal into oneself can lead (you) to depression.'

(29) *Le bonheur dépend de soi*
 'Happiness depends on oneself.'

To explain examples similar to (28)–(29), Zribi-Hertz (1990) suggests that *soi* comes in two variants: *soi* is either the non-nominative version of *on*, which does not need to be bound (it is a free arbitrary pronoun), or a variable that has to be bound by a universal quantifier. Instead of postulating homonymy, I propose that *soi* is a first-person oriented generic anaphor, and its twofold behavior is similar to that of other anaphors such as English *himself*, Icelandic *sig* or Mandarin *ziji* (see references in the introduction): in the general case, *soi* is locally bound (i.e. it obeys Condition A), but it can be exempt from locality requirements under logophoric conditions.⁶ What must now be understood is in which sense logophoricity makes acceptable examples like (26)–(29) where *soi* is not locally bound.

3.2 Logophoricity of *soi*

The notion of logophoricity has been originally coined for specific pronouns in West-African languages (e.g. in the Kwa language Ewe) that occur in environments such as indirect discourse and refer to the one “whose speech, thoughts, feelings, or general state of consciousness are reported” (Clements 1975: 141). As mentioned in the introduction, it has been observed that similarly, anaphors from

5. It could be argued that *soi* has a covert antecedent in (28) and (29), namely an arbitrary PRO subject of the noun *repli* ‘withdrawal’ or *bonheur* ‘happiness’. Zribi-Hertz (1990) mentions the following example with a free occurrence of *soi* (since *soi* cannot be bound by the specific antecedent Pierre).

(i) *Pierre essaie de trouver des idées utiles non seulement pour les autres, mais aussi pour soi.*
 ‘Peter tries to find ideas useful not only for others, but also for oneself.’

6. As will be explained in the conclusion, the exemption of *soi* is only apparent, since I will assume the existence of a logophoric operator locally binding the anaphor. I here adopt Charnavel and Sportiche’s (2016b) terminology, where ‘exempt’ is a descriptive term referring to anaphors that appear to be exempt from Condition A.

various languages can be exempt from Condition A when they are logophoric, i.e. when they occur in propositions expressing the perspective of their antecedent. As the notion of perspective center lacks precision, it has sometimes been proposed that perspective centers fall into different subtypes, which are relevant in different ways depending on languages. For instance, Sells (1987) distinguishes between Source (the one who is the intentional agent of the communication), Self (the one whose mental state or attitude the proposition describes) and Pivot (the one with respect to whose [time-space] location the content of the proposition is evaluated). To determine the type of logophoric center that is pertinent in the case of *soi*, I will here examine in turn the logophoric subtypes that have been shown to be relevant in French (Charnavel and Mateu 2015, Charnavel 2017), namely empathy loci, attitude holders and discourse participants.

According to Kuno (1987), the empathy locus is the event participant with whom the speaker empathizes, or identifies: a sentence containing an empathy locus describes the event from that empathy locus' point of view (even in non-attitude contexts). Kuno and Kaburaki (1977) use the Japanese giving verbs *yaru* and *kureru* as an illustration, which describe the event of giving from the perspective of the giver and the receiver, respectively; the reflexive *zibun*, when non-locally bound, must take the corresponding perspective center as antecedent (cf. Oshima 2006). The French anaphors *lui-même* and *son propre* are also sensitive to this type of perspective: they can be exempt from locality requirements when anteceded by an empathy locus (Charnavel 2017). But this cannot be the case of *soi*, since *soi* always involves empathy: as we have observed in the previous section, *soi* expresses a generalization based on the discourse participants' identification with the antecedent. Given that empathy is lexically encoded in *soi*, it cannot be relevant to distinguishing between exempt and non-exempt uses of *soi*.

The following examples reveal that the notion of attitude is not pertinent either for exempting *soi* from Condition A:

- (30) *Là-bas en Grèce, on_i garde confiance en soi_i* [= (20)]
 'Over there in Greece, one_i remains confident in oneself_i.'
- (31) **Là-bas en Grèce, on_i pense que l'Europe a peur de soi_i*
 '*Over there in Greece, one_i thinks that Europe fears oneself_i.'
- (32) *Ici en France, on_i pense que l'Europe a besoin de soi_i*
 'Here in France, one_i thinks that Europe needs oneself_i.'

In both (31) and (32), *on*, the antecedent of *soi*, is the subject of the attitude verb *pense* 'thinks' – i.e. the attitude holder – and *soi* occurs in its complement, namely in the attitude context whose center is its antecedent: in both cases, *soi* is in logophoric conditions related to attitude. Nevertheless, *soi* is only exempt from

Condition A in (32): *soi* is not acceptable in (31), where the antecedent is at the same distance from *soi* as in (32) (it is not in its minimal clause). This demonstrates that the attitudinal perspective is not a sufficient condition for exempting *soi* from Condition A (unlike German *man*, see Kratzer 1997).

The crucial difference between (31) and (32) is that the antecedent is speaker-exclusive in the former, but speaker-inclusive in the latter. Furthermore, Example (20) repeated above in (30) shows that a speaker-exclusive antecedent can antecede *soi* when it locally binds *soi*. This reveals that speaker inclusion in the antecedent is the relevant condition distinguishing between exempt and non-exempt *soi*.

The speaker's perspective thus creates the appropriate logophoric conditions exempting *soi* from Condition A. This requires the speaker to be included in the antecedent as in (32) vs. (31). This also requires that the speaker be the perspective center of the proposition containing *soi*. This explains the contrast between (32) and (33), which both involve speaker-inclusive *on* as antecedent of *soi*.

- (33) * *Ici en France, on_i entend dire par les Chinois que [l'Europe a besoin de soi_i]*
 '*Here in France, one_i hears from the Chinese that Europe needs oneself_i'

The bracketed clause containing *soi* expresses the antecedent's thoughts in (32), but not in (33), where it expresses the Chinese's. This is due to the fact that in the first case, *soi* appears in the complement of the attitude verb *pense* 'thinks' whose subject is the antecedent *on*, but in the second case, *soi* occurs in the complement of the verb *entend dire* 'hear' that shifts the perspective to the referent of the *by*-phrase.

The contrast between (32) and (31) on the one hand, and (33) on the other hand, thus demonstrates that *soi* does not have to be locally bound when its antecedent includes the speaker and is the perspective center of the clause containing it: it is the speaker's perspective that creates the relevant logophoric conditions exempting *soi* from Condition A. This is what makes Examples (32), (26)–(29), (8)–(9) and (11) acceptable (in the case of quantifiers like *personne* 'nobody' as antecedents, speaker inclusion means that the speaker is included in the domain of quantification).

Finally, the fact that the speaker's perspective plays a crucial role in the exemption of *soi* is corroborated by the existence of intervention effects. It has been shown that exempt anaphors are sensitive to perspective conflicts: for instance, in Mandarin, two clausemate long distance anaphors (*ziji*) have to corefer, and the pronouns *I/you (wo/ni)* trigger blocking effects (Huang and Liu 2001, a.o.). Similarly, (34) (from Sportiche et al. 2013: 435; cf. (6) too) shows that the

intervention of *tu* ‘you’ (vs. *la nation* ‘the nation’)⁷ gives rise to a conflict with the speaker’s perspective that is required to exempt *soi* from locality conditions.

- (34) *Aujourd’hui, on_i pense que {la nation/*tu} a(s) besoin de soi_i.*
 ‘Nowadays, one_i thinks that {the nation/*you} need(s) oneself_i.’

The contrast between (35) and (36) furthermore supports the hypothesis that the speaker’s perspective is only necessary when *soi* is not locally bound as in (35): *te* ‘you’ does not intervene in the case of locally bound *soi* in (36).⁸

- (35) **Si on_i pense que tu_s as besoin de soi_p...*
 ‘*When one_i thinks that you_s need oneself_p...’

- (36) *Si on_i te_s parle de soi_p...*
 ‘When one_i talks to you_s about oneself_p...’

4. Conclusion

In sum, the behavior of French *soi* supports the hypothesis that anaphors disobeying Condition A fall into one and the same category: they are exempt from locality requirements under logophoric conditions. Under closer scrutiny, it indeed turns

7. As noticed by a reviewer, the nation has to be one’s own nation here. This is because long distance *soi* imposes the speaker’s perspective.

8. The first-person pronoun *je* ‘I’ seems to trigger the same kind of intervention effects:
 [Context: the French President is talking:]

- (i) **En France, on_i aime sentir que j’ai besoin de soi_i.*
 ‘*In France, one_i likes feeling that I need oneself_i.’

Given our hypothesis about the speaker-oriented logophoricity of exempt *soi*, the ungrammaticality of (i) cannot be explained by perspective conflicts, since it is only about the speaker’s perspective here. But the unacceptability of (i) is due to a ban on partial overlap (similar to, though different from, Condition B, to which first and second-person French pronouns are not subject), which is independently observed in cases like (ii).

- (ii) **J’ai besoin de nous* (je ∈ nous)
 ‘*In France, I need us.’

Conversely, the first-person pronoun *me* does not intervene when *soi* is speaker-exclusive as in (iii) (from Rezac and Jouitteau 2008), since there is no overlap in reference (this can only be tested with locally bound *soi*, since non-locally bound *soi* is not acceptable when it is speaker-exclusive as we observed in (31)).

- (iii) *Partout où je vais, on_i me_s parle de soi_i et de sa vie* (me ∉ on)
 ‘Everywhere I go, one_i talks to me_s about oneself_i and one’s life.’

out that *soi* is not a long-distance anaphor that must be bound within its tensed clause as usually assumed, but it is a standard anaphor that has to be bound within its local domain (roughly, its minimal clause) except when the relevant logophoric conditions are met, in which case *soi* is not subject to any binding requirement. What obscures the behavior of *soi* is its specific generic properties: *soi* is a first-person oriented generic, which constrains its logophoric properties; that is, the relevant type of perspective exempting *soi* from Condition A is the speaker's perspective.

Thus, once the different factors at play have been disentangled, the case of French *soi* does not support the distinction between long distance anaphors and logophoric exempt anaphors, which means that the hypothesis of a crosslinguistically uniform binding domain remains viable (pending extensions to the other cases of so-called long-distance anaphors; see Charnavel and Sportiche 2016a, for the case of Icelandic *sig*).

It nevertheless remains to be the case that we have to distinguish between instances of anaphors obeying Condition A and instances of logophoric anaphors exempt from it. This also seems to go against ideals of parsimony, especially since the same anaphors exhibit the two types of behaviors: this is the case of *soi*, but also that of many other anaphors from various language families (English *himself*, Icelandic *sig*, Mandarin *ziji*, Japanese *zibun*...etc). Instead of postulating massive homonymy, I adopt the hypothesis that this twofold behavior can be accounted for in a parsimonious way if we suppose the existence of silent, syntactically represented logophoric operators, which locally bind the apparently exempt anaphors (and are coreferent with their antecedents), thereby determining their perspectival interpretation. In other words, exempt logophoric anaphors are no different from locally bound anaphors: both are subject to locality conditions, but they are fulfilled by covert antecedents in the former case, which only visibly affects their interpretation (see Charnavel 2017, for more details). That's why both types of anaphors have the same form in language after language: they are simply instances of the same lexical entry.

Acknowledgements

Thanks to Dominique Sportiche, two anonymous reviewers and the audience of LSRL46 for useful comments.

References

- Charnavel, Isabelle. 2017. "Exempt Anaphors and Logophoricity in French." In *Romance Languages and Linguistic Theory 12. Selected papers from the 45th Linguistic Symposium on Romance Languages (LSRL)*, ed. by Ruth E. V. Lopes, Juanito Ornelas de Avelar, and Sonia M. L. Cyrino, 15–28. Amsterdam: John Benjamins Publishing Company.
- Charnavel, Isabelle, and Victoria Mateu. 2015. "The Clitic Binding Restriction revisited: Antilogophoricity in Clitic Clusters." *The Linguistic Review* 32 (4): 671–701. <https://doi.org/10.1515/tlr-2015-0007>
- Charnavel, Isabelle, and Dominique Sportiche. 2016a. "Unifying Long Distance Binding: Icelandic *sig* is Clause-bound, or Logophoric." In *Proceedings of the 42th annual meeting of the Berkeley Linguistics Society (BLS 42)*, ed. by Emily Clem, Virginia Dawson, Alice Shen, Amalia Horan Skilton, Geoff Bacon, Andrew Cheng, and Erik Hans Maier, 47–62. Berkeley: Berkeley Linguistics Society.
- Charnavel, Isabelle, and Dominique Sportiche. 2016b. "Anaphor Binding – What French Inanimate Anaphors Show." *Linguistic Inquiry* 47 (1): 35–87. https://doi.org/10.1162/LING_a_00204
- Chierchia, Gennaro. 1998. "Reference to Kinds across Language." *Natural Language Semantics* 6 (4): 339–405. <https://doi.org/10.1023/A:1008324218506>
- Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origins, and Use*. New York: Praeger.
- Clements, George N. 1975. "The Logophoric Pronoun in Ewe: Its Role in Discourse." *Journal of West African Languages* 10: 141–177.
- Cole, Peter, Gabriella Hermon, and C. -T. James Huang. 2006. "Long-Distance Binding in Asian Languages." In *The Blackwell Companion to Syntax*, ed. by Martin Everaert, and Henk Van Riemsdijk, 21–84. Malden, MA: Blackwell. <https://doi.org/10.1002/9780470996591.ch39>
- Creissels, Denis. 2008. "Impersonal Pronouns and Coreference: The Case of French *on*." Ms. University of Lyon.
- Huang, C. -T. James, and C. -S. Luther Liu. 2001. "Logophoricity, Attitudes, and *ziji* at the Interface." In *Syntax and Semantics 33: Long Distance Reflexives*, ed. by Peter Cole, Gabriella Hermon, and C. -T. James Huang, 141–195. San Diego: Academic Press.
- Kayne, Richard S. 2003. "Person Morphemes and Reflexives in Italian, French and related Languages." In *The Syntax of Italian Dialects*, 102–136. New York: Oxford University Press.
- Kratzer, Angelika. 1997. "German Impersonal Pronouns and Logophoricity." Talk given at Sinn und Bedeutung 2, Berlin.
- Kuno, Susumu. 1987. *Functional Syntax: Anaphora, Discourse and Empathy*. Chicago: University of Chicago Press.
- Kuno, Susumu, and Etsuko Kaburaki. 1977. "Empathy and Syntax." *Linguistic Inquiry*: 627–672.
- Malamud, Sophia. 2012. "Impersonal indexicals: *one*, *you*, *man*, and *du*." *The Journal of Comparative Germanic Linguistics* 15 (1): 1–48. <https://doi.org/10.1007/s10828-012-9047-6>
- Maling, Joan. 1984. "Non-clause bounded reflexives in Modern Icelandic." *Linguistic and Philosophy* 7: 211–241. <https://doi.org/10.1007/BF00627706>
- Manzini, M. Rita, and Kenneth Wexler. 1987. "Parameters, Binding Theory, and Learnability." *Linguistic Inquiry*: 413–444.
- Moltmann, Frederick. 2006. "Generic *one*, Arbitrary PRO, and the First Person." *Natural Language Semantics* 14 (3): 257–281. <https://doi.org/10.1007/s11050-006-9002-7>

- Oshima, David Y. 2006. "Perspectives in Reported Discourse." Ph.D. Dissertation, Stanford University.
- Pearson, Hazel. 2013. "A Judge-free Semantics for Predicates of Personal Taste". *Journal of Semantics* 30 (1): 103–154. <https://doi.org/10.1093/jos/ff5001>
- Pica, Pierre. 1984. "On the Distinction between Argumental and Non-Argumental Anaphors." In *Sentential Complementation*, ed. by Wim de Geest, and Yvan Putseys, 185–193. Dordrecht: Foris.
- Pica, Pierre. 1987. "On the Nature of the Reflexivization Cycle." In *Proceedings of the North East Linguistic Society* 17 (2): 483–500.
- Pica, Pierre. 1991. "On the Interaction between Antecedent-Government and Binding: The Case of Long-Distance Reflexivization." In *Long Distance Anaphora*, ed. by Jan Koster, and Eric Reuland, 119–135. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511627835.006>
- Pollard, Carl, and Ivan A. Sag. 1992. "Anaphors in English and the Scope of Binding Theory". *Linguistic Inquiry* 23: 261–303.
- Reinhart, Tanya, and Eric Reuland. 1993. "Reflexivity." *Linguistic Inquiry* 24: 657–720.
- Reuland, Eric. 2006. "Icelandic Logophoric Anaphora." In *The Blackwell Companion to Syntax*, ed. by Martin Everaert, and Henk Van Riemsdijk, 544–557. Malden, MA: Blackwell. <https://doi.org/10.1002/9780470996591.ch33>
- Rezac, Milan, and Mélanie Jouitteau. 2008. "The Nature of French 'Impersonal' Indexicals *tu* and *on*." Talk given at the Workshop on Impersonal Pronouns, Paris.
- Ronat, Mitsou. 1982. "Une solution pour un apparent contre-exemple à la théorie du liage [a Solution for an Apparent Counter-example to the Binding Theory]." *Lingvisticae Investigationes* 6 (1): 189–196. <https://doi.org/10.1075/li.6.1.07ron>
- Sells, Peter. 1987. "Aspects of Logophoricity." *Linguistic Inquiry* 18: 445–479.
- Sportiche, Dominique, Hilda Koopman, and Edward Stabler. 2013. *An Introduction to Syntactic Analysis and Theory*. Malden, MA: Wiley-Blackwell.
- Zribi-Hertz, Anne. 1990. "*Lui-même* argument et le concept de « pronom A » [*Lui-même* Argument and the Concept of 'A Pronoun']". *Langages* 97: 100–127. <https://doi.org/10.3406/lgge.1990.1576>

French negative concord and discord

An experimental investigation of contextual and prosodic disambiguation

Viviane Déprez and Jeremy Yeaton

Rutgers University

Evidence that DN readings arise in solid NC languages more than previously thought (Déprez et al. 2015) underscore the importance of investigating the factors governing their emergence to deepen our understanding of Negative Concord. This paper examines the roles of context and prosody in disambiguating simple French transitive sentences with two NCIs (*personne*, *rien*) experimentally. French speakers were recorded reading target sentences with NCIs in NC or DN eliciting contexts followed by a verification statement judged as T/F and analyzed as context-matching or not. Context-matching targets were extracted using Praat and analyzed using ProsodyPro (Xu 2013). Sentence contour, average pitch and F0 peak value on NCIs were compared across conditions. Our results provide solid evidence that in French, context influence interpretation and NC and DN interpretations are prosodically distinguished.

Keywords: negative concord, double negation, prosody of French negative concord, context influence

1. Introduction

French, sequences of potentially negative expressions like *personne* and *rien*, referred to here as Negative Concord Items (NCI) (Watanabe 2004),¹ can have two possible interpretations; a sentence like (1a) can have a Negative Concord (NC) reading as in (1b) or a Double Negation (DN) reading as in (1c), in which two negations cancel each other out to produce a positive reading.

1. The expressions that participate in negative concord dependencies are also often referred to as n-words. We chose to avoid this term here to avoid any unpleasant connotation.

- (1) a. *Personne ne dit rien.*
- b. *Nobody says anything* = Everyone is silent = NC
- c. *Nobody says nothing* = Everybody talks. = DN

On a par with other Romance languages, French is generally regarded as a concord language (Zeijlstra 2004, De Swart 2010). As is known, however, double negation readings of NCI sequences are possible and not uncommon (Déprez 2000, Corblin and Tovena 2003). Still quite unexplored, however, are the different factors that govern the choice between NC and DN interpretations for French speakers. The question of whether NC is a default reading in French as well as the influence of the morpho-syntactic nature of NCI, and their structural position on this choice was examined in previous experimental work (Déprez et al. 2013, Déprez 2014). Processing considerations influencing interpretation are discussed in Corblin (1996). The present study focuses on the role of context, and especially of prosody on the choice of NC/DN interpretation in standard French. Quite a few authors have noted the importance of prosody in influencing the interpretation of French NCI sequences (Corblin and Tovena 2003, Corblin et al. 2004) but up to now, there has been no experimental investigation of how and to what extent NC and DN could be prosodically distinguished. Here we report the result of a production experiment designed to start answering this question.

2. Theoretical background

A popular theoretical stand in generative approaches to Negative Concord has been to consider that there is a macro-parameter that divides NC languages from DN languages (Zanuttini 1991, Haegeman 1995, Zeijlstra 2004, a.o.). In main European languages, the partition would by and large run along the Romance vs Germanic divide, though with noted exceptions. The predictions of such views are that no real NC/DN ambiguities should be observed in languages, and hence, the emergence of NC in DN languages or DN in NC languages is often cast as marked or anomalous and, as such, not considered as part of what the grammars allow. There have been essentially 3 families of theoretical accounts of the phenomenon of Negative Concord, which differ according to the status given to NCIs. In the first type of account, NCI are considered to be non-negative expressions, of either existential or universal nature, and the single negation reading of concord comes from a single sentential negation operator that is overt (Giannakidou 2000) or covert (Zeijlstra 2004). Here, the predictions are that DN readings in sequences of NCI should be no more available than with sequences of negative polarity items like *anything* or *anyone*. In the second type of account, NCI are negative expressions

and concord readings are obtained through a semantic process of Resumptive Quantification, a subtype of polyadic quantification (May 1990, De Swart and Sag 2002, Déprez 2007, 2000). Here, in contrast, sequences of NCI are predicted to be ambiguous and can in principle lead to either NC or DN readings, depending on whether they are interpreted under scopal interaction or resumptive quantification. Left unaccounted on this view, however, is how this choice operates and thus, how languages could differ in their NC/DN distribution. Finally, in a third type of approach, NCI are taken to be ambiguous expressions that are sometimes negative, sometimes not. The ambiguity has been regarded as lexical by some researchers (Herburger 2001) and as structural by others, shaped by the internal syntax of NCIs (Déprez 2000). Déprez (2011) argued that structural considerations affect the interpretability of the negative features of NCIs both within the DP and in the sentence in which they occur. Very briefly, negative features become interpretable at phase edge, but how they get to these edges, and how phases are established in a language can be subject to variation. The prediction of this approach is that NC and DN readings are subject to morpho-syntactic and structural conditions that can vary both language-internally and cross-linguistically, though the principles of negative interpretability remain unchanged.

3. Experimental background

A previous experimental study on French negative dependencies used a picture to meaning matching choice task without verbal contexts to probe NC and DN interpretations (Déprez et al. 2013, Déprez 2014, Déprez et al. 2015). Contrasting with the expectations fostered by much of the literature on French, the results showed no solid preference for NC. French speakers preferred NC only slightly but not significantly in NCI sequences with pronominal expressions like *personne* and *rien* (noted *Pro* in Figure 1), but significantly preferred DN readings when NCI sequence featured full DP expressions like *aucun enfant* (noted *DP* in Figure 1) in subject positions, although NC readings were not excluded in these cases either. The graph in Figure 1 sums up the central results of this first experimental investigation.

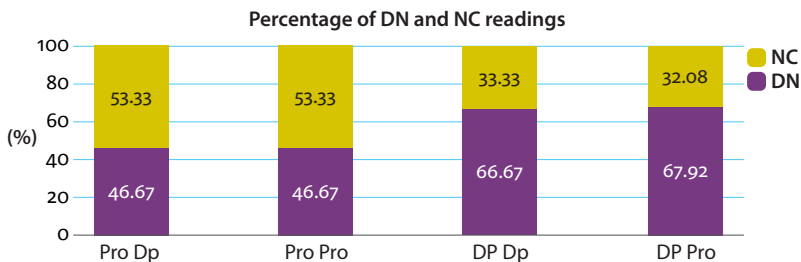


Figure 1. Morphosyntactic factors influencing NC vs DN readings

We see that the choice of NC or DN is almost equivalent when the subject is a pronominal NCI like *personne*: NC at 53% and DN at 46%. So these structures are clearly ambiguous. But when the NCI is a full DP like *aucun enfant* in subject position, then the DN readings are significantly preferred at 66%, $p < 0.001$. These results confirm that morpho-syntactic and structural factors can strongly influence the NC/DN interpretation choice within a single language. Yet since structures with pronominal NCI were shown to be ambiguous between NC and DN readings, further factors need to be considered for these particular cases, and more specifically, context and prosody. Concerning the influence of context on interpretation, it has been quite generally assumed that in order to be felicitous, a DN sentence requires a special context. Yet the specific nature of this context is not always made clear. Many have noted that the contradiction of a previous negation facilitated these readings. More specifically, as proposed in Puskás (2012), DN sentences are said to be felicitous when they negate a discourse old negation with a discourse new negation.² This context type, however, is said to facilitate DN readings in all types of languages, independently of whether or not they have been regarded as NC or DN languages. In French, however, contexts of this type are not necessary to elicit a potential DN reading for NCI sequences, even if, of course, they remain possible (Corblin and Tovena 2003, Déprez 2000). So a sentence like (2) can be understood with a DN reading. Assume a context where a restaurant bill is split in a group and everyone must contribute.

- (2) *Tout le monde a bien réglé sa part ? Oui en effet, personne n'a rien payé.*
 'Did everyone pay their part? Yes, indeed. Nobody paid nothing.'

2. That denial of a previous negation is not always necessary for DN even for supposed DN languages like English is shown by examples like (i), said by a resolute mother determined to protect her child:

- (i) There is nothing I would not do for my child.
 (American Crime, commercial for an ABC TV show).

In this example, DN is used to strengthen an affirmation.

Here, the context facilitates DN because the positive meaning of the double negation sentence (everyone paid something) is the expected outcome. Since we were interested specifically in how this ambiguity is resolved in French, and whether DN readings can be elicited in contexts that do not feature a contradiction, these are the types of context that were used in our design.

The literature reveals some previous experimental research on NC/DN prosody, though at this point, none that we are aware of for French. Prosody has first been investigated for Afrikaans in a production study by Huddleston (2010), who found NC and DN to be distinguished. A similar study on Dutch by Fonville (2013) did not produce such clear results. It is notable, however, that in the experimental design of both of these production experiments, the interpretation of the actual target sentences was presupposed to match the one the context intended to facilitate. As there was no meaning verification test, however, it is ultimately unclear what interpretation the speakers really had. Some of the observed fluctuation may well result from this design problem, so that in the end, claims of a prosodic differentiation or lack thereof are difficult to evaluate. Other prosodic studies on Catalan (Espinal et al. 2011, Espinal et al. 2016), found a clear NC/DN distinction. This study, however, looked at the meaning of single NCI in answers to negative questions, which proved to be ambiguous in Catalan, but are not felt to be so in French, where they have been observed to systematically lead to a DN interpretation (Corblin and Tovená 2003). In view of our previous experimental results and these preceding prosody studies, the study reported here sought to answer the following research questions: 1. Does context succeed in influencing/determining NC/DN interpretation in French? 2. What is the role of prosody? Does French manifest a particular mapping between prosody and interpretation in NCI sequences? If so which one?

4. Experimental design

To investigate these questions, we designed a production experiment. Participants were asked to read simple ambiguous transitive sentences in context and were recorded. The test sentences all followed the form [*personne* Verb *rien* PP], and were embedded in short contexts facilitating a NC or DN interpretation, as illustrated in (3) and (4) below. Importantly, our DN facilitating contexts did not contain negative propositions, which have been argued to facilitate DN in all languages, even strict NC ones (Puskás 2012), but rather were simple statements that described situations compatible with a DN reading for our targets. Following the reading, subjects were presented with a verification statement that served to check whether participant interpretation indeed matched the one intended by the context. For

example, respective contexts were as in (3) for the ambiguous target sentence: “*personne ne boit rien dans les soirées*” (Nobody drinks anything/nothing at parties),

(3) NC Context:

Dans notre famille, on est tous allergique à l'alcool :

‘In our family, we are all allergic to alcohol.’

Personne ne boit rien dans les soirées.

‘Nobody drinks anything at parties.’

(4) DN Context:

Chez les jeunes, la consommation d'alcool est effrayante :

‘In the youth population, alcohol consumption is frightening.’

Personne ne boit rien dans les soirées.

‘Nobody drinks nothing at parties.’

For both (3) and (4) above, the verification statement was as in (5) below.

(5) *Ils ne boivent pas d'alcool.*

‘They don’t drink alcohol.’

In the NC context presented in (3), the interpretation intended would yield a response of “True” to the verification statement: since everyone is allergic to alcohol, no person in the family drinks any alcohol, i.e.: no alcohol is consumed, therefore “they don’t drink alcohol” would be true. By contrast, the sentence would be false following the DN eliciting context in (4). Since the consumption of alcohol among the youth has reached frightening levels, one understands that at parties, no youth fails to drink, therefore everyone is drinking. Hence the verification statement, “they don’t drink alcohol” would be false.

There were 5 categories of stimuli, each with 8 items per category: (1) ambiguous criticals in DN contexts, as in (4) above; (2) ambiguous criticals in NC contexts, as in (3) above; (3) single negation controls with a single NCI in subject position (*Personne ne boit ça*, ‘nobody drinks this’); (4) single negation controls with an NCI in object position (*Les gens ne boivent rien*, ‘People drink nothing’), and (5) fillers (*Les enfants boivent du jus*, ‘Children drink juice’). The verification statements were balanced 4T/4F within each category. With 40 items per participant and 28 participants, 1120 total responses were collected. In order to avoid ordering effects, the items were pseudo-randomized in blocks. Each block featured one item per category, randomized within the block, and the blocks were appended one after the other also in random order such that every participant saw all 40 items. In order to control for prosodic factors, the same number of syllables was used in each of the target sentences. Two syllables for the subject, one syllable for the verb, and one syllable for the object. In order to avoid having the sentence boundary tone fall on the object NCI, each sentence was concluded by a

prepositional phrase of at least two syllables. In addition, the same 8 monosyllabic verbs, of comparable frequency, were used in each category.

5. Participants and procedure

28 native French speakers (18F, 10M, age 18–45), most students at the University of Lyon were recorded at the L2C2 Bron, France, on an Asus Orion PRO gaming headset with a noise filtering microphone. Participants were asked to read contexts and targets first silently, and then aloud, as if telling a story to a child. The recording was self-paced. After they finished recording, participants answered the verification question T/F. Participants were first presented with two practice items and the total experimental time took at most 20 minutes.

6. Analysis

Two main analyses were conducted: contextual interpretation and acoustic/prosodic. For the contextual interpretation analysis, T/F responses were coded as +/-context matching, as well as NC/DN. For the acoustic/ prosodic analysis, the critical sentences (*personne ne* [verb] *rien*) [PP]) were excised from context using Audacity 2.0.6 and time aligned, matching phonemes and syllables to waveform in Praat (Boersma and Weenink 2015) using EasyAlign (Goldman 2011). A second Praat script called ProsodyPro (Xu 2013) was used to extract F0 values.

7. Results: Context influence

In control contexts, interpretations correctly matching the context (NC match = NC interpretation in NC context, DN match = DN interpretation in DN contexts) were provided at 96.21%, from which it can be safely concluded that the task was well understood by the participants. For critical items, overall interpretation matched context at almost 80%, though a bit more in NC context (87%) and a bit less in DN contexts (72%). This clearly shows that context was successful in determining interpretation and is thus a strong factor that influences the NC/DN choice in French.

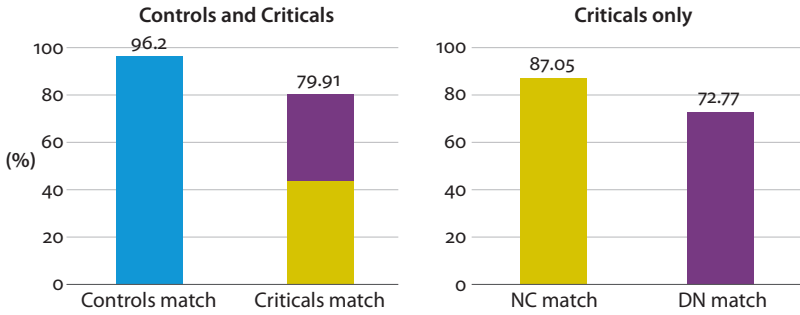


Figure 2. Matching responses in Controls and Critical items

When the contexts results are examined by participant, it can be seen that NCI sequences were not equally ambiguous for all participants; some (7/28) strongly favored NC readings in both NC (Bule black) and DN contexts (= NC mismatch blue grey), and 1/28 strongly favored DN readings in all DN (yellow stripes) and NC contexts (= DN mismatch purple stripes). But for the majority of participants (20/28), sequences of NCIs items are clearly ambiguous: they favor NC interpretations in NC contexts and DN interpretations in DN contexts for the same sentences. Our prosodic analysis was restricted to these 20 participants and their 278 matching items (138 DN, 140 NC). Hence the context verification control in our experiment allowed us to confirm the ambiguity, and it also allowed us to filter out from the prosodic analysis the speakers and productions that did not have the intended meaning: participants and items mismatching the intended meanings were removed from further analysis to ensure that the productions we looked at clearly had the meaning we were interested in.

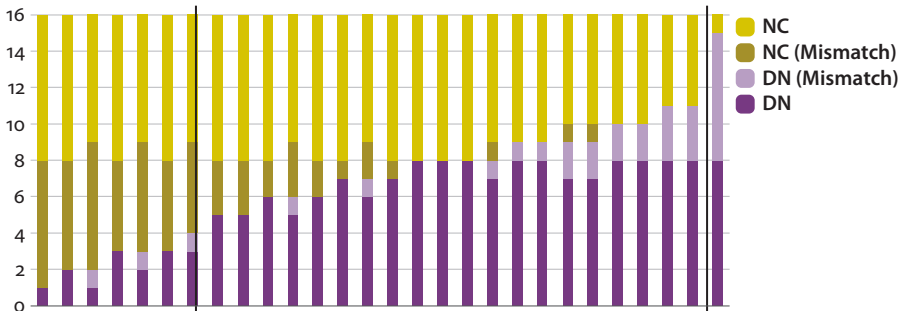


Figure 3. Matching and non-matching responses per participants.

8. Results: Acoustic analysis

From the 278 recordings that were kept for analysis after meaning verification, we excised our disambiguated target sentence using the open source free audio editor software Audacity 2. 1.2:

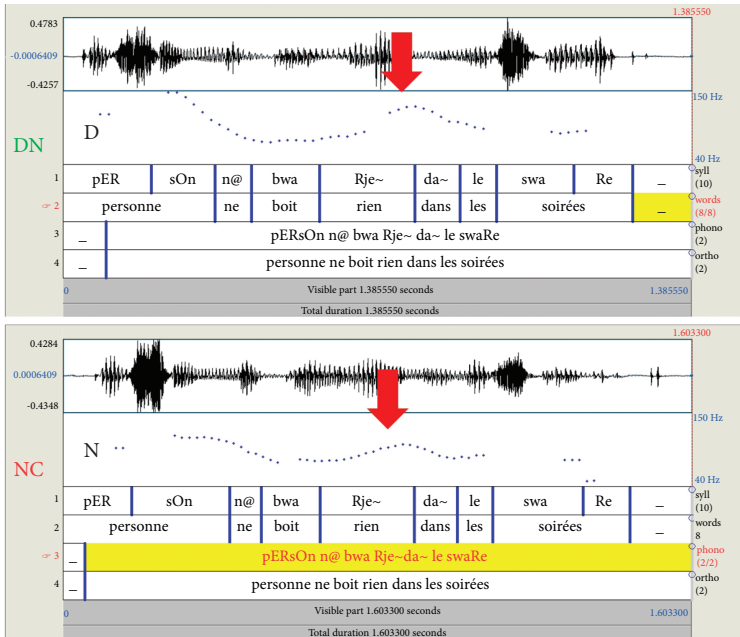


Figure 4. Sample F0 contour

Before we turn to the results of our acoustic analysis, let us first look at a sample F0 contour produced by Praat. Note the difference seen in these contours between the realization of *rien* in the DN reading as opposed to the NC reading. The DN reading here has a higher F0 peak on this item than the NC one. For each syllable in the sentence, ProsodyPro (Xu 2013) was used to Extract (10) time-normalized F0 values. The time-normalized values helped give an idea of how pitch was changing over the course of the syllable as opposed to over the course of time. Thus, instead of each segment having the same period of time (e.g.: 0.1 seconds), each segment represented 0.1 syllables. These ten values were then aligned by syllable across all 8 utterances in a given condition for a given participant. They were then averaged together to create an “average contour” for a participant in a given condition. These participant contours were, in turn, further averaged together to give an idea for the overall behavior in a condition. These were weighted by number

of utterances used, since we only used context-matched recordings, and plotted to see the overall F0 shape of the speakers in a given condition:

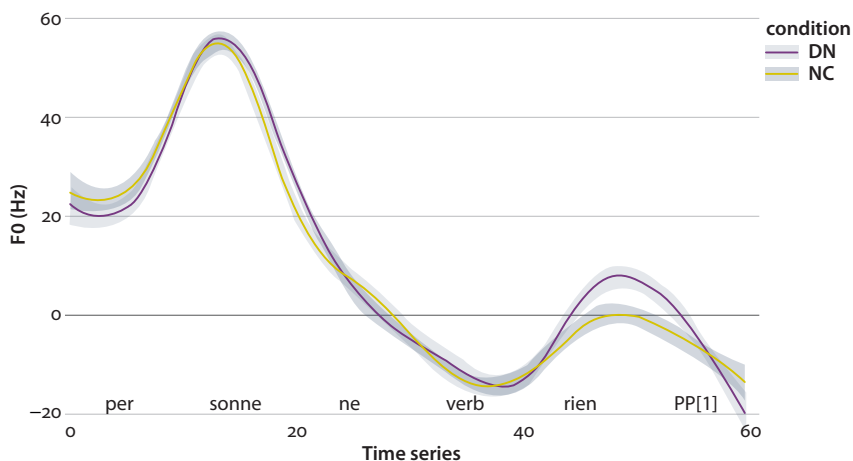


Figure 5. De-meanned averaged contour per condition

Figure 5 shows the average of the de-meanned F0 values for all DN matching and all NC matching items respectively. It can be noted that pitch is higher and rises and falls more sharply on the object *rien* in the DN condition as compared to the NC condition. The difference in the average pitch of *rien* in the DN and NC condition for all subjects combined is significant at the $p < 0.05$ level. The peak F0 values for the NCI *rien* for all of the participants in the two conditions (DN = 16.90Hz vs NC = 6.54Hz) is also significantly different ($p < 0.05$). This is illustrated in Figure 6 both quantitatively on the right and graphically on the left.

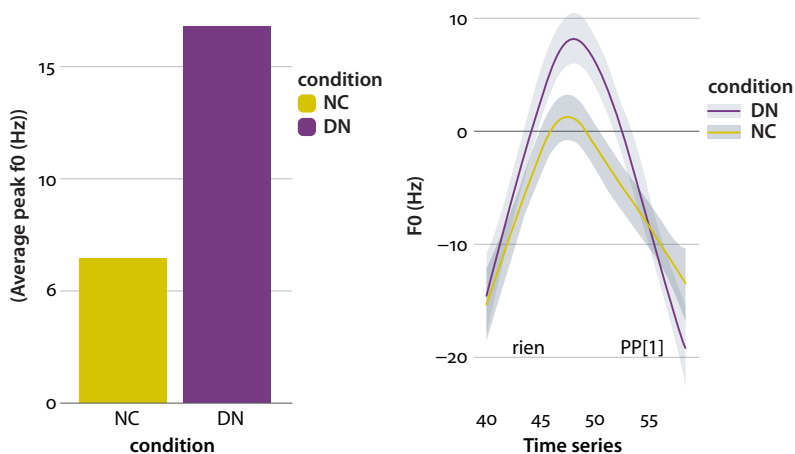


Figure 6. Peak distinction on *rien*

In addition to the peak values, the difference between the peak value on *rien* and the lowest pitch of the following syllable, which represents the pitch drop after *rien* was also significant between conditions, as can be seen below in Figure 7. The bar graph on the right represents the average pitch difference between the highest peak on *rien* and the lowest pitch on the following syllable (=pitch drop) greater for DN than for NC which is graphically represented on the left.

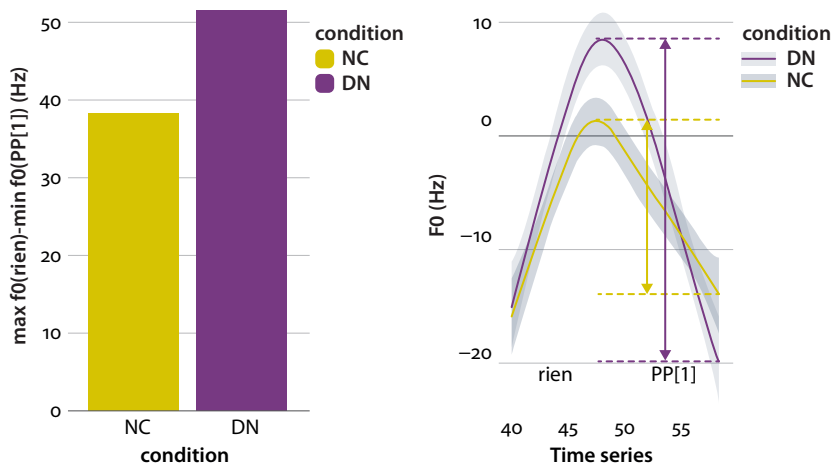


Figure 7. Pitch drop after *rien*

The average F0 contours of all of the participants were then compared and two main patterns were found. The first and less common pattern (4 participants), shown below in Figure 8, is characterized by a higher peak value on the second syllable of the subject NCI *personne* in the DN condition, with little difference on the object NCI. This pattern corresponds to the one suggested in Corblin et al. (2003) as typical for DN readings. For Corblin et al., emphasis on the subject NCI is what is assumed to be distinctive for DN. It is interesting to note that this pattern, although indeed represented in our data, turned out to be used by only a small minority of our participants.

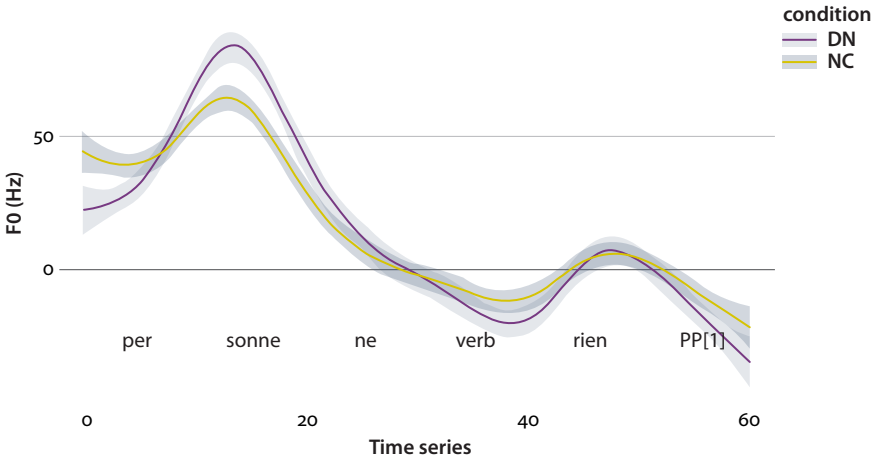


Figure 8. Subject NCI emphasis, DN pattern

The other pattern, which was exhibited by 14 of our participants, featured no discernible difference on the NCI first (*personne*) but a much higher peak value on second NCI (*rien*) in the DN condition.

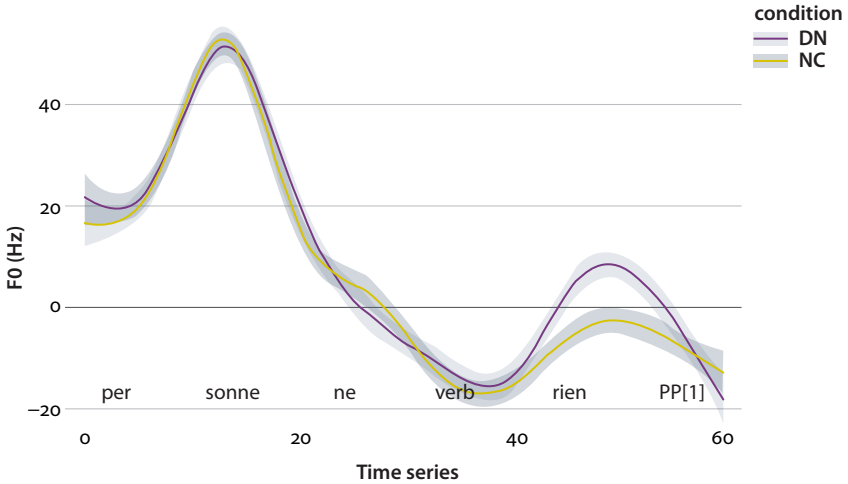


Figure 9. Object NCI emphasis, DN pattern

Two of our participants did not conform to either of these main patterns. Neither did they have a common pattern with one another.

9. Discussion

The acoustic measures described above clearly indicate that speakers make a distinction in production between the NC and DN readings. However, these measures do not yet constitute a prosodic analysis of this distinction, since a prosodic analysis is based on differences that can be perceived by speakers. We think, however, that the acoustic differences uncovered in this experiment are indeed indicative of a prosodic difference and we suggest that the NC and DN readings are characterized by a phrasing difference as well as by a difference in tone. For NC, we suggest, the prosodic analysis, given in 1 below, where *personne* forms its own accentual phrase, with a low pitch accent on the first syllable, and a high boundary tone marking the subject accentual phrase on the second syllable. The verb and the NCI object *rien* are then phrased together with a low pitch accent on the V and a high boundary tone on *rien* marking the VP accentual right edge phrase boundary. Finally, the prepositional phrase closes the sentence with a phrase final low boundary tone.

- (6) NC: Focus on *personne*; *rien* is phrased as part of VP

$$\begin{array}{ccccccc} L^* & H- & & L^* & H- & & L\% \\ (([_{DP}Personne]_{AP}) & ([_{VP}ne \text{ Verb } rien]_{AP}) & .. & ([_{PP}...PP...]_{AP}) &]_{IP} \end{array}$$

For the DN reading in contrast, we suggest that *rien* forms its own accentual phrase, separate from that of the verb, which is possibly de-phrased (Féry 2001) or phrased along with the subject (Avanzi et al. 2014). Furthermore, because *rien* is an accentual phrase by itself, it carries a Low+High or rising boundary tone.

- (7) DN: Focus on *personne*; V is ‘dephrased’ (Féry 2001); Focus on *rien* which forms its own phonological phrase.

$$\begin{array}{ccccccc} L^* & H- & & L^* & LH- & & L\% \\ (([_{DP}Personne]_{AP}) & ne \text{ Verb } ([_{DP}rien]_{AP}) & ([_{PP}...PP...]_{AP}) &]_{IP} \end{array}$$

Following Féry (2001), we suggest that this phrasing indicates that in DN, the second NCI is under focus, which is here manifested by an increased F0 peak and followed by a more abrupt fall on the next syllable, which could mark the end of the *rien* phrase.

Support for this analysis is provided by the Prosograms (Mertens, to appear) we obtained for our target sentences, of which a representative sample contrasting realization of the DN reading (above) and the NC reading (under) of a particular speaker is provided in Figure 10.

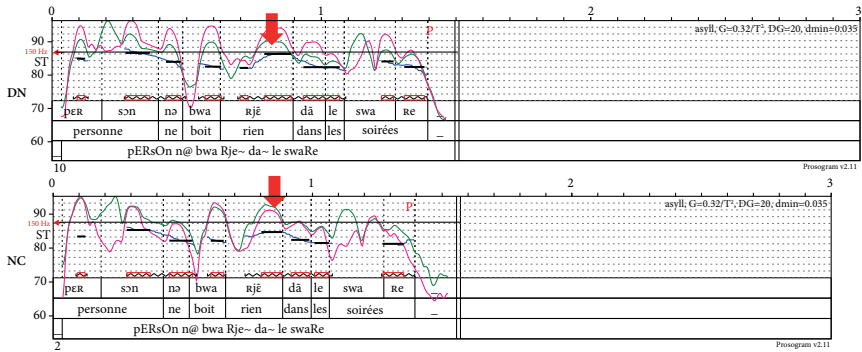


Figure 10. Sample contrastive Prosograms

Prosogram, designed by Piet Mertens, is a software for the semi-automatic transcription of prosody. It takes a time aligned phonetic annotation (here delivered by Praat) and outputs a representation that is an estimation of the pitch contour that can be perceived by an average listener. This representation is a stylization of the F0 curve of vowel nuclei, which aims at the reconstruction of the perceived pitch contour, based on a psycho-acoustic model of tonal perception.

In the sample provided in Figure 8, we see the difference between the stylized perceived contour of the DN and NC production of the target sentence *personne n'aime rien*. Two facts are notable in these representations. First note that the DN clearly has a higher tone on *rien* than the NC one. As you can see the thick bar is set at 150hz in both diagrams. In DN, the tone on *rien* reaches the bar, while in NC, in contrast, it remains clearly below. Furthermore, for DN, the representation shows a Low-High tone on the monosyllabic *rien*, which we take to be representative of *rien* forming a phrase by itself. In contrast, we see a flatter and lower high tone on *rien* in the NC reading. This is a distinction that matches our proposed prosodic analysis and suggests that the difference is indeed perceivable by an average French speaker.

10. Conclusions

First, our results confirm that both the NC and the DN readings are available for most French speakers, and that this ambiguity is not marginal. These results do not quite fit a NC/DN cross-linguistic parametric divide since, on such a view, French would be unclassifiable. Hence, an empirically correct account of French negative dependencies requires moving away from a macro-parametric divide. We have also seen that context quite generally succeeds in influencing the interpretation, although not equally for all speakers. At this point, an interesting question remains

as to whether this speaker distinction reflects a distinction in their grammars or in their capacities to compute or process ambiguities. Finally, we have shown that, in production, the NC and DN interpretations are acoustically distinguished by a statistically significant higher F0 peak on the second NCI, and a steeper fall after it in the DN reading. We have further proposed that this acoustic difference signals a prosodic difference in phrasing and in tone. We argued that in DN, but not in NC, the second NCI can form its own prosodic phrase independent from the VP and bear a Low-High boundary tone. We would like to further suggest that, possibly, the two prosodic phrasing for NC and DN correspond to two distinct syntactic structures; One for NC, where the NCI remains or reconstructs inside the VP, so that its negative features are not interpretable on Déprez (2011)'s feature interpretability condition, which proposes that negative features are interpretable only when they occur at phase edges (DP,vP,CP); and the other for DN where, under focus, *rien* is interpreted at vP edge, so that its negative feature can be interpreted at the sentential and thus, at the propositional level.

References

- Avanzi, Mathieu, George Christodoulides, and Elisabeth Delais-Roussarie. 2014. "Prosodic Phrasing of SVO Sentences in French." In *Social and Linguistic Speech Prosody. Proceedings of the 7th international conference on Speech Prosody*, ed. by Nick Campbell, Dafydd Gibbon, and Daniel Hirst, 703–707. Dublin: Science Foundation Ireland.
- Boersma, Paul, and David Weenink. 2015. *Praat: Doing phonetics by computer* [Computer program]. Version 6.0.07. Retrieved 30 May 2015 from <http://www.praat.org>
- Corblin, Francis. 1996. "Multiple negation processing in natural language." *Theoria. A Swedish Journal of Philosophy* 62 (3): 214–260.
- Corblin, Francis, Viviane Déprez, Henriette De Swart, and Lucia Tovenà. 2004. "Negative Concord." In *Handbook of French semantics*, ed. by Henriette De Swart and Corblin, 417–455. Stanford: CSLI.
- Corblin, Francis, and Lucia Tovenà. 2003. "L'expression de la négation dans les langues romanes." In *Les langues romanes: Problèmes de la phrase simple*, ed. by Daniele Godard, 281–343. Paris: CNRS Editions.
- De Swart, Henriette. 2010. *Expressing and Interpreting Negation: An OT Typology*. Dordrecht: Springer. <https://doi.org/10.1007/978-90-481-3162-4>
- De Swart, Henriette, and Ivan A. Sag. 2002. "Negation and negative concord in Romance." *Linguistics and Philosophy* 25: 373–417. <https://doi.org/10.1023/A:1020823106639>
- Déprez, Viviane. 2000. "Parallel (A)symmetries and the Internal Structure of Negative Expressions." *Natural Language and Linguistic Theory* 18 (2): 253–342. <https://doi.org/10.1023/A:1006449808181>
- Déprez, Viviane. 2007. "On the structuring role of grammaticalized morpho-syntactic features." In *Proceeding of NELS 37*, ed. by Emily Elfner, and Martin Walkow, 31–44. Amherst, MA: GLSA.

- Déprez, Viviane. 2011. "Atoms of negation: An outside-in micro-parametric approach to negative concord." In *The Evolution of Negation: Beyond the Jespersen Cycle*, ed. by Pierre Larrivée, and Richard Ingham, 221–272. Amsterdam: John Benjamins.
<https://doi.org/10.1515/9783110238617.221>
- Déprez, Viviane. 2014. "NC vs. DN: a macroparametric variation? What can an experimental investigation tell us?" Keynote paper presented at the Linguistic Symposium of Romance Languages 44, University of Western Ontario.
- Déprez, Viviane, Anne Cheyls, and Pierre Larrivée. 2013. "When and How is Concord preferred? An experimental approach." In *L'Interface langage-cognition The Language-cognition Interface. Actes du 19e Congrès International des Linguistes, Genève, 22–27 juillet 2013*, ed. by Stephen R. Anderson, Jacques Moeschler, and Fabienne Reboul. Genève: Droz.
- Déprez, Viviane, Susagna Tubau, Anne Cheylus, and M. Teresa Espinal. 2015. "Double Negation in a Negative Concord language: An experimental investigation." *Lingua* 163: 75–107.
<https://doi.org/10.1016/j.lingua.2015.05.012>
- Espinal, M. Teresa, and Pilar Prieto. 2011. "Intonational encoding of double negation in Catalan." *Journal of Pragmatics* 43: 2392–2410. <https://doi.org/10.1016/j.pragma.2011.03.002>
- Espinal, M. Teresa, Susagna Tubau, Joan Borràs-Comes, and Pilar Prieto. 2016. "Double negation in Catalan and Spanish. Interaction between syntax and prosody". In *Negation and polarity: Experimental perspectives*, ed. by Pierre Larrivée, and Chungmin Lee, 145–176. Berlin: Springer. https://doi.org/10.1007/978-3-319-17464-8_7
- Féry, Caroline. 2001. "Focus and Phrasing in French." In *Audiatu Vox Sapientiae. A Festschrift for Arnim von Stechow*, ed. by Caroline Féry, and Wolfgang Sternefeld, 153–181. Berlin: Akademie-Verlag.
- Fonville, Renata. 2013. "The role of intonation in the use of double negatives in Dutch." Master's thesis, Utrecht University.
- Giannakidou, Anastasia. 2000. "Negative ... concord." *Natural language and linguistic theory* 18: 457–523. <https://doi.org/10.1023/A:1006477315705>
- Goldman, Jean-Philippe. 2011. "EasyAlign: an automatic phonetic alignment tool under Praat." In *Proceedings of the 12th Annual Conference of the International Speech Communication Association 2011 (InterSpeech 2011)*, 3240–3242. Red Hook, NY: Curran Associates.
- Haegeman, Liliane. 1995. *The syntax of negation*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9780511519727>
- Herburger, Elena. 2001. "The negative concord puzzle revisited." *Natural Language Semantics* 9: 289–333. <https://doi.org/10.1023/A:1014205526722>
- Huddleston, Kathleen M. 2010. "Negative Indefinites in Afrikaans." Ph.D. dissertation, Utrecht University.
- May, Robert. 1990. "A note on quantifier absorption." *The Linguistic Review* 7: 121–127.
<https://doi.org/10.1515/tlir.1990.7.1.121>
- Mertens, Piet. To appear. "The Prosogram model for pitch stylization and its applications in intonation transcription." In *Prosodic Theory and Practice*, ed. by Jonathan A. Barnes, and Stefanie Shattuck-Hufnagel. Cambridge: MIT Press.
- Puskás, Genoveva. 2012. "Licensing double negation in NC and non-NC languages." *Natural Language and Linguistic Theory* 30: 611–649. <https://doi.org/10.1007/s11049-011-9163-z>
- Watanabe, Akira. 2004. "The genesis of negative concord: Syntax and morphology of negative doubling." *Linguistic Inquiry* 35 (4): 559–612. <https://doi.org/10.1162/0024389042350497>

- Xu, Yi. 2013. "ProsodyPro – A Tool for Large-scale Systematic Prosody Analysis." In *Proceedings of Tools and Resources for the Analysis of Speech Prosody (TRASP 2013)*, ed. by Brigitte Bigi, and Daniel Hirst, 7–10. Aix-en-Provence: Laboratoire Parole et Langage (LPL).
- Zanuttini, Rafaella. 1991. "Syntactic properties of sentential negation." Ph.D. dissertation, University of Pennsylvania.
- Zeijlstra, Hedde. 2004. "Sentential Negation and Negative Concord." Ph.D. dissertation, University of Amsterdam.

Dimensions of variation

The inflected construction in the dialect of Delia (Caltanissetta)

Vincenzo Nicolò Di Caro and Giuliana Giusti
Ca' Foscari University of Venice

This paper provides an overview of the different dimensions of variation found in the Inflected Construction (IC) (cf. Cardinaletti and Giusti 2001, 2003) arising with motion verbs in most Sicilian dialects. It does so by discussing data from the dialect of Delia (Caltanissetta) at face value with the dialect of Marsala (Trapani). The micro-variation emerging from the discussion challenges Cruschina's (2013) morphomic account; while it is easily captured by Cardinaletti and Giusti's (to appear) hypothesis that a semi-functional V1 fills a remerged functional head (*t*-T, *c*-C). Merger in *t* or *c* is not only restricted to lexically specified forms of V1 but, in given feature combinations, also to lexically specified forms of V2 in T or C.

Keywords: inflected construction, pseudo-coordination, sicilian dialects, cartographic syntax, motion verbs, restructuring verbs

1. Introduction

Cardinaletti and Giusti (henceforth, C&G) study a pseudo-coordination construction in the dialect of Marsala (in the province of Trapani), which they call the Inflected Construction (henceforth, IC) (C&G 2001, 2003).¹ They show that, parallel to pseudo-coordinations previously studied in English and Swedish, the IC in Marsalese features the following components: (i) a verb (V1), usually of motion,

1. Previous literature on Italo-Romance has labelled this construction in different ways. Traditional literature emphasises its paratactic origin (cf. Ascoli 1896, 1901, Sorrento 1950, Rohlfs 1969, Stefanini 1970, Leone 1973, 1978 and Sornicola 1976); Cruschina (2013) calls it the Doubly Inflected Construction; Ledgeway (2015) treats it as one of different grammaticalized pseudo-coordinations in the dialects of Apulia.

taken from a restricted class of restructuring verbs; (ii) a sometimes optional connecting element *a*; (iii) a lexical verb (V2), sharing mood, tense and person features with V1.

C&G (2001, 2003) report the following restrictions in the dialect spoken in Marsala (Trapani):

- i. The IC is only possible in the 1/2/3SG and 3PL of the indicative present and in the 2SG of the imperative;
- ii. The only possible V1s are the most basic motion verbs ‘go’ and ‘come’, together with ‘come by’ and the motion causative verb ‘send’.

C&G (2001: 383–5) also note that the IC in Marsalese can display invariant forms of ‘go’ as V1, similarly to what happens with auxiliary verbs ‘have’ and ‘stay’ (cf. (1)–(3)), where the invariant form is allowed with the three singular persons of the present indicative and is incompatible with the plural persons and with tenses other than present indicative:

- (1) a. *Vaju / va a ppigghju u pani.*
go.1SG / go a fetch.1SG the bread
‘I go and fetch the bread.’
b. *Vannu / va a ppigghjanu u pani.*
go.3PL / go a fetch.3PL the bread
‘They go and fetch the bread.’
- (2) a. *Un ci haju / ha statu mai.*
not there.CL have.1SG / have been never
‘I have never been there.’
b. *Un ci hai / ha statu mai.*
not there.CL have.2SG / have been never
‘You have never been there.’
- (3) a. *Ci staju / sta ennu.*
there.CL stay.1SG / stay going
‘I’m going there.’
b. *Ci stai / sta ennu.*
there.CL stay.2SG / stay going
‘You’re going there.’

C&G’s (2001: 397–407) account for these restrictions proposing that V1 is a semi-lexical verb, namely a lexical verb that loses its argument structure, being merged as a functional head. In a parametric perspective, they claim that V1 in the IC is merged at the point where the language realizes the inflected V2, which is subject to parametric variation: in Marsalese it is merged in T, like in Italian and unlike

English or Swedish (where it is merged lower). From this high vs low merging position of the semi-lexical verb, they derive a number of macro-parametric differences between Marsalese and the Germanic languages, which cannot be dealt with here due to space reasons.

C&G propose that the lexicon specifies which forms of given verbs can fill the V1 position of the IC, which generally appear to be “less rich” in inflection.² C&G therefore predict that it is possible to find variation on single verbal forms that can enter the IC as V1, as is in fact the case. Their framework does not make any prediction on possible restrictions on V2, which in Marsalese can be any lexical verb compatible with andative aspect.

In this paper, we concentrate on the dialect of Delia (in the province of Caltanissetta). This dialect presents some interesting points of micro-variation, confirming C&G’s general framework and leading us to a more precise formulation of the syntactic structure of the IC and to positing further lexical restrictions on V2.

The structure of the paper is as follows. Section 2 presents subsequent literature on Sicilian and other Southern-Italian varieties and briefly discusses possible competing hypotheses. Section 3 presents the morphology of the IC in Deliano and shows that there are different lexical restrictions on V1 and further restrictions on V2. Section 4 provides a formal structural representation of the IC in Deliano. Section 5 deals with the interaction of the IC with Cinque’s (2006) hierarchy of restructuring verbs. Section 6 draws the conclusions.

2. Subsequent literature on the inflected construction

Manzini and Savoia (2005: 688–701) provide a very wide overview of the phenomenon and show that in Southern Italian dialects, V1 can be filled by other verbs that behave as restructuring verbs in the Italian infinitival construction. The IC in Sicilian, Apulian, and partially Calabrian dialects, can display other moods, tenses and persons but is crucially always limited to simple forms. Manzini and Savoia (2005: 689) also show that verbs other than ‘go’ (notably ‘stay’ but also ‘come’) can display invariant forms. This kind of variation is predicted by C&G (2001, 2003).

2. In recent terms, this can be defined as a “nano-parameter” in the sense of Biberauer and Roberts (2012: 268), who propose that parametric variation displays different degrees of variation. (i) Macro-parameters regulate all elements of a given type. (ii) Meso-parameters regulate a featurally specifiable subset of the elements of a given type. (iii) Micro-parameters regulate the smallest definable sub-class of elements of a given type. (iv) Nano-parameters regulate one or more individual lexical items; this is the case of the Sicilian IC, as further argued in this paper.

Cruschina (2013) observes that the inflectional restrictions described by C&G (2001, 2003) for Marsalese correspond to Maiden's (2004) N-pattern (see also Dressler and Thornton 1991, Thornton 2007) and proposes to treat the IC as a Serial Verb Construction (cf. Aikhenvald and Dixon 2006), which in some varieties but not in others, is restricted to this morphomic pattern. Cruschina's purely morphomic analysis however does not predict the following facts, that will be exemplified in this paper on the basis of Deliano: (1) the restriction to simple tenses of the paradigm even in the most liberal varieties; (2) the restrictions to what restructuring verbs can fill V1; (3) the different morphomic patterns for different tenses;³ (4) the different behaviour of invariant forms in different dialects.

There is no space in this paper to do justice to the theory of morphemics, arising from Aronoff's (1994) hypothesis of the autonomy of morphology. For what is relevant to our discussion, it is sufficient to know that it gives autonomous value to given patterns created by the inflectional morphology of a given lexical item. Thus, if a given variety attributes a given pattern to the IC, it is not predicted that different V1s or different combinations of V1 and V2 should display different patterns.

Ledgeway (2015) concentrates on the dialects of Apulia, where the IC is attested with 'go', 'stand', and 'want' as V1s. In particular, 'stand' and 'go' display different degrees of reduction of the inflected forms, throughout the region, while 'want' never does. According to Ledgeway, different degrees of reduction correspond to different degrees of grammaticalization of V1. For reasons of space, we cannot deal with the Apulian cases, thus referring to Cardinaletti and Giusti (to appear) who note that there is no difference in behavior of invariant vs. inflected forms with respect to clitic climbing, which can safely be taken as a clear diagnostics of monoclausality. We will see that Deliano does not generally allow for invariant forms except for one single combination of V1-V2 in the imperative.

3. The morphology of the inflected construction in Deliano

The IC in Deliano is possible with the four V1s, namely 'go', 'come', 'come by' and 'send', found in Marsalese (cf. C&G 2001, 2003) and with a fifth V1, *accuminciari* 'start'. These five verbs do not display a homogenous behavior in their morphological distribution.

With the four motion verbs, the IC follows the N-pattern, as in Marsalese; namely, it is restricted to 1/2/3SG and 3PL in the indicative present and to the 2SG in the imperative:

3. See Di Caro and Giusti (2015) for a discussion of the facts in (1)–(3) which Cruschina's (2013) analysis does not take into consideration.

- (4) a. *Vaju / vjignu / passu / mannu a ppigliu lu pani*
 go.1SG / come.1SG / come-by.1SG / send.1SG a fetch.1SG the bread
 ‘I go / come / come by/ send somebody to fetch the bread.’
- b. **Jammu / *vinjimmu / *passammu / *mannammu a ppigliammu lu*
 go.1PL / come.1PL / come-by.1PL / send.1PL a fetch.1PL the
 pani.
 bread
 ‘We go / come / come by/ send somebody to fetch the bread.’
- (5) a. *Va / vjini / passa / manna (a)*
 go.IMPR.2SG / come.IMPR.2SG / come-by.IMPR.2SG / send.IMPR.2SG a
ppiglia lu pani!
 fetch.IMPR.2SG the bread
 ‘Go / come / come by/ send somebody to fetch the bread!’
- b. **Jiti / *viniti / *passati / *mannati a*
 go.IMPR.2PL / come.IMPR.2PL / come-by.IMPR.2PL / send.IMPR.2PL a
ppigliati lu pani!
 fetch.IMPR.2PL the bread
 ‘Go / come / come by/ send somebody to fetch the bread!’

The inceptive verb *accuminciari* is instead limited to 1SG and 3PL of the indicative present:

- (6) a. *Accuminciu a ffazzu la spisa.*
 start.1SG a do.1SG the shopping
 ‘I start doing the shopping.’
- b. *Accumincianu a ffannu la spisa.*
 start.3PL a do.3PL the shopping
 ‘They start doing the shopping.’
- c. **Accumincia a ffa la spisa.*
 start.3SG a do.3SG the shopping
 ‘(S)he starts doing the shopping.’

In a morphomic perspective, one is led to note that the restrictions in (6) are reminiscent of Maiden’s (2004) U-pattern, with the proviso that Deliano does not have subjunctive (which is instead part of the U-pattern). Cruschina’s hypothesis that the IC is associated to the N-pattern in Marsalese should be reformulated for Deliano assuming the N-pattern with motion verbs and the U-pattern with the inceptive verb. This weakens the predictive power of a framework that attributes independent predictive power to given patterns.

Unlike Marsalese, the IC in Deliano is also possible in the indicative preterite with the five V1s already cited and a crucial restriction on V2. Interestingly, the

preterite displays a third pattern, which is not present in the morphomic literature, with ungrammatical 2SG/PL:

- (7) a. *Jivu a ffici la spisa du voti.*
 go.PAST.1SG *a* do.PAST.1SG the shopping two times
 ‘I went to do the shopping twice.’
- b. **Jisti a ffacisti la spisa du voti.*
 go.PAST.2SG *a* do.PAST.2SG the shopping two times
 ‘You went to do the shopping twice.’
- c. *Ji a ffici la spisa du voti.*
 go.PAST.3SG *a* do.PAST.3SG the shopping two times
 ‘(S)he went to do the shopping twice.’
- d. *Jammu a fficimu la spisa du voti.*
 go.PAST.1PL *a* do.PAST.1PL the shopping two times
 ‘We went to do the shopping twice.’
- e. **Jistivu a ffacistivu la spisa du voti.*
 go.PAST.2PL *a* do.PAST.2PL the shopping two times
 ‘You went to do the shopping twice.’
- f. *Jiru a fficiru la spisa du voti.*
 go.PAST.3PL *a* do.PAST.3PL the shopping two times
 ‘They went to do the shopping twice.’

In Di Caro and Giusti (2015: 410–11), we call this the W-pattern for purely descriptive reasons, attributing no explicatory value to this label.

Di Caro (2015: 50) observes that the possible V2 forms in the preterite are all rhizotonic forms of verbs that come from Latin third conjugation:⁴

- (8) a. *fari* (‘do’ or ‘make’) < *facĕre*
 b. *vìviri* (‘drink’) < *bibĕre*
 c. *scrìviri* (‘write’) < *scribĕre*
 d. *mìntiri* (‘put’) < *mittĕre*
 e. *diri* (‘say’ or ‘tell’) < *dicĕre*
 f. *pèrdiri* (‘lose’) < *perdĕre*
 g. *chjùiri* (‘shut’) < *cludĕre*
 h. *‘nchjùiri* (‘shut in’) < *includĕre*

These verbs are different from the regular first conjugation verbs, such as *pigliari* ‘fetch’, which display six arhizotonic forms (e.g. *pigliàvu*, *pigliàsti*, *piglià*, *pigliàm-mu*, *pigliàstivu*, *pigliàru*).

4. The verb *vidiri* ‘see’, although diachronically derived from a verb of Latin second conjugation, i.e. *vidĕre*, displays the same inflection of the verbs listed in (8).

In fact, two additional lexical verbs appear in the indicative preterite, which do not derive from Latin third conjugation, but synchronically pattern as irregular verbs of the first conjugation, like *fari* ‘do’ above; namely, *stari* ‘stay’ and *dari* ‘give’. The entire paradigm of these two verbs is made of rhizotonic forms, but the W-pattern is still manifested, as shown in (9):

- (9) a. *Ci jivu a ddetti nna manu d’ajutu.*
to-him go.PAST.1SG a give.PAST.1SG a hand of help
‘I went to give him a hand.’
- b. **Ci jisti a ddasti nna manu d’ajutu.*
to-him go.PAST.2SG a give.PAST.2SG a hand of help
‘You went to give him a hand.’
- c. *Ci ji a ddetti nna manu d’ajutu.*
to-him go.PAST.3SG a give.PAST.3SG a hand of help
‘(S)he went to give him a hand.’
- d. *Ci jammu a ddjittimu nna manu d’ajutu.*
to-him go.PAST.1PL a give.PAST.1PL a hand of help
‘We went to give him a hand.’
- e. **Ci jistivu a ddàstivu nna manu d’ajutu.*
to-him go.PAST.2PL a give.PAST.2PL a hand of help
‘You went to give him a hand.’
- f. *Ci jiru a ddjittiru nna manu d’ajutu.*
to-him go.PAST.3PL a give.PAST.3PL a hand of help
‘They went to give him a hand.’

In this case, the W-pattern correlates with the forms displaying apophony (the root vowel turns from *-a-* to *-e-* / *-ji-*). It must be observed that there are verbs, like *sapiri* ‘know’, featuring preterite rhizotonic forms but no apophony (i.e. *-a-* is the root vowel in all the paradigm) for 1/3SG, 1/3PL, which do not display the IC in the preterite at all:

- (10) a. **Vinni a ssappi di sta novità.*
come.PAST.1SG a know.PAST.1SG of this news
- a’. *Vinni a ssappiri di sta novità.*
come.PAST.1SG to know.INF of this news
‘I was made aware of this news.’
- b. **Vinnimu a ssàppimu di sta novità.*
come.PAST.1PL a know.PAST.1PL of this news
- b’. *Vinnimu a ssappiri di sta novità.*
come.PAST.1PL to know.INF of this news
‘We were made aware of this news.’

It is not clear to us how syllable structure and apophony may correlate with this. C&G (2001, 2003) reduce all restrictions to lexical information on V1. In their perspective, it is not expected that phonology be directly relevant to these restrictions.

Finally, a further verb can occur only in the preterite, namely *arristari* ‘remain’, it also follows the W-pattern, but only in a fixed combination with ‘give’ as V2, which is interpreted as ‘still owing something to somebody’, as in (11):

- (11) a. *Ci arristavu a ddetti deci euro.*
to-him remain.PAST.1SG a give.PAST.1SG ten euro
‘I still owe him ten euro(s).’
- b. **Ci arristasti a ddasti deci euro.*
to-him remain.PAST.2SG a give.PAST.2SG ten euro
‘You still owe him ten euro(s).’
- c. *Ci arristà a ddetti deci euro.*
to-him remain.PAST.3SG a give.PAST.3SG ten euro
‘(S)he still owes him ten euro(s).’
- d. *Ci arristammu a ddjittimu deci euro.*
to-him remain.PAST.1PL a give.PAST.1PL ten euro
‘We still owe him ten euro(s).’
- e. **Ci arristàstivu a ddàstivu deci euro.*
to-him remain.PAST.2PL a give.PAST.2PL ten euro
‘You still owe him ten euro(s).’
- f. *Ci arristaru a ddjittiru deci euro.*
to-him remain.PAST.3PL a give.PAST.3PL ten euro
‘They still owe him ten euro(s).’

Note that *arristari* in the IC cannot be interpreted as having its literal meaning of ‘remaining in the same place’. With this interpretation, it can only select the infinitive, as in (12b):

- (12) a. **M' arristavu a bbitti la partita nn'iddru*
REFL remain.PAST.1SG a see.PAST.1SG the match at him
- b. *M' arristavu a bbidiri la partita nn'iddru.*
REFL remain.PAST.1SG to see.INF the match at him
‘I stayed to watch the match at his place.’

Note that even in the metaphorical meaning, *arristari* cannot enter the IC in the indicative present (13) or in the imperative (14):

- (13) a. **Ci arrjistu a ddugnu sempri sordi a ma frati.*
to-him remain.1SG a give.1SG always money to my brother
- b. *Ci arrjistu a ddari sempri sordi a ma frati.*
to-him remain.1SG to give.INF always money to my brother
‘I always owe my brother some money.’

- (14) a. **Tu arrèstacci a ddùna sordi, accusi s' inzigna.*
 you remain.IMPR.2SG-to-him a give.IMPR.2SG money so REFL
 teach.3SG
- b. *Tu arrèstacci a ddari sordi, accusi s' inzigna.*
 you remain.IMPR.2SG-to-him to give.INF money so REFL teach.3SG
 'Don't pay him back completely, that will teach him!'

We can detect an increasing degree of lexical restriction in the examples provided so far, which cannot even be captured by a hierarchy among the patterns or tenses. Some lexical items enter the IC in one pattern in the present, some in others in other patterns in the present and preterite, other still appear only in the preterite in a combination of V1 and V2. This appears to be a step towards idiomaticization reminiscent to the more extreme idiomatic status of Italian *vattelappesca* (lit. 'go and fish it!') and of regional Italian *vattelaccerca* (lit. 'go and look for it!'), which both have the meaning of 'goodness knows' and can be reconstructed as imperative forms.

These intricacies can be captured by reformulating C&G's proposal so that the lexical restrictions are not only on V1 but also on V2. A morphomic account would have to assume that different combinations of V1 and V2 enter the IC according to different patterns, for no obvious reasons.

4. The structural representation of the IC

C&G (to appear) suggest that, in the IC, V1 is merged in *t*, a functional head immediately higher than T parasitically checking its Tense and Agreement on T, as in (15):

- (15) $[_{tP} t [a [_{TP} T [_{VP} V \dots]]]]$

In Deliano, when *t* and T share [PRES] features and 1SG and 3PL agreement, V1 is limited to 'go', 'come', 'come by', 'send', and 'start', as in (16a); when *t* and T share [PRES] features and 2SG and 3SG agreement, V1 is limited to 'go', 'come', 'come by', 'send' and cannot include 'start', as in (16b). In both cases, V2 can be any lexical verb semantically compatible with the andative semantics:

- (16) a. $[_{tP} [_{V1} \text{go/come/come by/send/start}] + t_{[PRES; 1SG/3PL]} [a [_{TP} V2 + T_{[PRES; 1SG/3PL]}]]]$
- b. $[_{tP} [_{V1} \text{go/come/come by/send}] + t_{[PRES; 2/3SG]} [a [_{TP} V2 + T_{[PRES; 2/3SG]}]]]$

With different person specifications, the construction is ungrammatical. Deliano can express the same andative aspect with a restructured infinitive, whose structure, we assume, is the same as in Italian (cf. Cinque 2006, Cardinaletti and Shlonsky 2004).

When *t* and T share [PAST] features and 1/3SG, 1/3PL agreement, V1 is still limited to ‘go’, ‘come’, ‘come by’, ‘send’, and ‘start’, as in the present, but V2 is restricted to the verbs listed in (8) and ‘start’, as in (17):

- (17) [_{tP} [_{V1} go/come/come by/send/start]+*t*_[PAST; 1/3SG/1/3PL] [_a [_{TP} [_{V2} do/drink/write/put/tell/lose/shut/stay]+T_[PAST; 1/3SG/1/3PL] [_{VP} V2...]]]]

Insertion of ‘give’ as V2 extends the range of V1 to ‘remain’, as in (18):

- (18) [_{tP} [_{V1} go/come/come by/send/start/remain]+*t*_[PAST; 1/3SG/1/3PL] [_a [_{TP} [_{V2} give]+T_[PAST; 1/3SG/1/3PL] [_{VP} V2...]]]]

C&G (to appear) do not provide the structure of the IC in the imperative but C&G (2001: 399) claim that the imperative is merged in a higher functional head in the complementizer layer, parallel to Jussive proposed by Zanuttini (2008: 196). Thus, the imperative should generate a complex *cP*-CPsystem, as in (19), which *mutatis mutandis* behaves like the *tP*-TP in (15) above:

- (19) [_{cP} V1 [(*a*) [_{CP} V2 [_{TP} V2 [_{VP} V2]]]]]

The presence of the connecting element *a* depends on the selection of V1. The variation in (20) is therefore expected. In (20a) ‘go’ and ‘come’ in the imperative do not take the connecting element. In (20b) ‘come by’ and ‘send’ must display the connecting element, as shown by the obligatory syntactic doubling on the following syllable:

- (20) a. *Va* / *Vjini* (**a*) *piglia* *lu pani!*
go.IMPR.2SG / come.IMPR.2SG *a* fetch.IMPR.2SG the bread
‘Go / Come and fetch the bread!’
- b. *Passa* / *Manna* *a ppiglia* *lu pani!*
come-by.IMPR.2SG / send.IMPR.2SG *a* fetch.IMPR.2SG the bread
‘Come by / Send somebody to fetch the bread!’

Structure (19) has therefore two realizations, as in (21):

- (21) a. [_{cP} [_{V1} go/come]+*c*_[JUSS; 2SG] [_{CP} V2+*C*_[JUSS; 2SG] [TP ...
b. [_{cP} [_{V1} come by/send]+*c*_[JUSS; 2SG] [_a [_{CP} V2+*C*_[JUSS; 2SG] [TP ...

The two realizations in (21) can co-occur, with a fixed order, namely the one given in (22a), where ‘go/come’ precede ‘come by/send’, as exemplified by grammati-

cal (23). The logically possible reversed order given in (22b) gives ungrammatical results, as exemplified in (24):

- (22) a. [_{CP} [_{V1} go/come] + c_[JUSS; 2SG] [_{CP} [_{V1} come by/send] + c_[JUSS; 2SG] [*a* [_{CP} V2 + C_[JUSS; 2SG] [TP ...
- b. [_{CP} [_{V1} come by/send] + c_[JUSS; 2SG] [*a* [_{CP} [_{V1} go/come] + c_[JUSS; 2SG] [_{CP} V2 + C_[JUSS; 2SG] [TP ...
- (23) a. *Vjini /Va manna a ppiglia lu libbru!*
 come/go.IMPR.2SG send.IMPR.2SG a fetch.IMPR.2SG the book
 ‘Come / Go and send somebody to fetch the book!’
- b. *Vjini /Va passa a appiglia lu libbru!*
 come/go.IMPR.2SG come-by.IMPR.2SG a fetch.IMPR.2SG the book
 ‘Go and come by to fetch the book!’
- (24) a. **Manna a bbjini piglia lu libbru!*
 send.IMPR.2SG a come.IMPR.2SG fetch.IMPR.2SG the book
 ‘Send somebody to come and fetch the book!’
- b. **Passa a bba piglia lu libbru!*
 come-by.IMPR.2SG a go.IMPR.2SG fetch.IMPR.2SG the book
 ‘Come by and go to fetch the book!’

Recursion is not possible in the indicative where all verbs select the connecting element *a*, as shown in (25):

- (25) a. **Vjignu a mmanu a ppigliu lu libbru.*
 come.1SG a send.1SG a fetch.1SG the book
 ‘I come and send somebody to fetch the book.’
- b. **Vaju a ppassu a ppigliu lu libbru.*
 go.1SG a come-by.1SG a fetch.1SG the book
 ‘I go to come by and fetch the book.’

The possibility of recursion in the imperative is therefore due to the two different constructions with and without *a* emerging with the two different V1s in the imperative.

Note that the imperative displays the only uninflected form found in the IC in Deliano, namely *va* ‘go’ which can however only be combined with a 2PL imperative ‘go’, as in (26a). No other V2 is possible with *va*, as in (26b). Nor is the 2SG possible, as shown in (26c):

- (26) a. *Va jitivinni!*
 go.IMPR.2SG go.IMPR.2PL-REFL.CL-LOC.CL
 ‘Go away!’

- b. **Va travagliati!*
 go.IMPR.2SG work.IMPR.2PL
 ‘Go to work!’
- c. **Va vatinni!*
 go.IMPR.2SG go.IMPR.2SG REFL.CL-LOC.CL
 ‘Go away!’

Once again, this kind of information can only be captured by a parameter that regards the specific lexical item.

5. Interaction with the cartographic hierarchy

According to Cinque’s (2006: 90) cartographic hierarchy of clausal functional projections, the andative functional head is rather low: lower than inceptive and causative, as in (27a). This is confirmed by the infinitival construction in Deliano, as in (27b):

- (27) a. Causative > Inceptive > Andative > V
 b. *Lu fazzu accuminciari a gghjiri a ppigliari lu pani sulu.*
 him.CL make.1SG start.INF to go.INF to fetch.INF the bread alone
 ‘I’ll make him start going to fetch the bread on his own.’

In (28), violation of the hierarchy (27a) results in ungrammaticality:

- (28) a. **Lu vaju a ffari pigliari a ma ma’.*
 it.CL go.1SG to make.INF fetch.INF to my mum
 ‘I’ll go to have my mother fetch it.’
 b. **L’ accuminciu a ffari pigliari a ma ma’.*
 it.CL start.1SG to make.INF fetch.INF to my mum
 ‘I’ll start having my mother fetch it.’

We now observe how the verbs that enter the IC as V1 interact with such a hierarchy. C&G (2001, 2003) have already shown that in Marsalese the motion verbs in V1 are not part of the hierarchy, as expected in the analysis in (27) above. In fact, the hierarchy regulates the portion of structure below TP and does not regard V1 in *t* in (15)–(18) or in *c* in (19) and (21). This derives the fact that in Marsalese the motion verb in V1 can co-occur with the low andative head.

Deliano confirms this prediction. Not only the motion verb as V1 in (29a) but also the inceptive ‘start’ as V1 in (29b) can precede the highest restructuring verb considered here, namely causative ‘make’:

- (29) a. *Lu vaju a ffazzu jiri a ppigliari a ma ma'.*
 it.CL go.1SG a make.1SG go.INF to fetch.INF to my mum
 'I'm going to have my mother go and fetch it.'
- b. *L' accuminciu a ffazzu jiri a ppigliari a ma ma'.*
 it.CL start.1SG a make.1SG go.INF to fetch.INF to my mum
 'I'll start having my mother go and fetch it.'

Since Deliano displays motion and inceptive verbs as V1, it allows us a further observation. There is a hierarchy of insertion of V1 in the IC, which works in the opposite direction of Cinque's hierarchy. In (30) we see that motion must precede inceptive in the IC, unlike what we observe in (31), where inceptive precedes motion in the infinitival construction:

- (30) a. *Lu vaju a accuminciu a ffari.*
 it.CL go.1SG a start.1SG to make.INF
 'I'll go and start doing it.'
- b. **L' accuminciu a bbaju a ffari.*
 it.CL start.1SG a go.1SG to make.INF
 'I'll start going to do it.'
- (31) a. **Lu vaju a accuminciari a ffari.*
 it.CL go.1SG to start.INF to make.INF
 'I'll go and start doing it.'
- b. *L' accuminciu a gghjiri a ffari.*
 it.CL start.1SG to go.INF to make.INF
 'I'll start going to do it.'

In C&G's framework the exact point of insertion of a semi-lexical verb as V1 in the IC is independent of Cinque's (2006) hierarchy because it depends on the position in which V+T is realized in that given language. Thus, V1 is expected to immediately precede V2, and it is V2 to be the highest verb in Cinque's functional hierarchy and, as such, to be merged with the clausal T-features. This leaves the ungrammaticality of (30b) unexplained, which may be due to the fact that *accuminciari* is not fully productive in the indicative present, while 'come/go' is. Note however that such ungrammaticality would not be accounted for in any other competing analysis. We leave this small point to further research.

6. Conclusions

In this paper, we have presented a fine-grained description of the IC in Deliano, which displays interesting differences as well as substantial similarities to what has been reported for Marsalese by C&G's (2001, 2003) original work.

The variation observed in Deliano confirms C&G's analysis along the following lines. The micro-variation in the different forms (tense, mood, and persons) that can occur as V1 in the IC supports C&G's hypothesis that the lexicon of a language must specify what individual forms of the paradigm of a given verb can enter the IC as V1. This would give rise not only to the already noted N-pattern and U-pattern, but also to a novel W-pattern displayed in the preterite. The preterite has also brought up a new piece of evidence to assume that the IC is strictly related to lexical specification also as regards V2.

We have provided a formalization of what needs to be assumed in the lexicon of Deliano, showing that C&G's proposal of analysing the IC as the possible insertion of V1 in a dedicated functional head checking verbal and nominal features parasitically onto the canonical functional head hosting V2 is correct.

Moreover, we have confirmed in Deliano C&G's observation that the complex syntactic structure of the IC does not respect the cartographic hierarchy. We have also found out that there is a hierarchy in the IC that even goes in the opposite direction, namely when an inceptive verb co-occurs with a motion verb in the IC, inceptive must follow motion, while motion follows inceptive in restructured infinitives.

Acknowledgements

This paper was presented at the 42nd Incontro di Grammatica Generativa, held in Lecce at Università del Salento on February 19, 2016, and at the 46th Linguistic Symposium on Romance Languages, held in New York at Stony Brook University on April 2, 2016. We wish to thank Andrea Calabrese, Roberta D'Alessandro, Ion Giurgea, Rita Manzini, Henk Van Riemsdijk, Ian Roberts, and two anonymous reviewers for stimulating remarks. All errors remain our own. Although the paper has been written four hands at all stages, for the sake of the Italian academic law, Giuliana Giusti is responsible for Sections 1–2 and 4, while Vincenzo Nicolò Di Caro is responsible for Sections 3 and 5–6.

References

- Aikhenvald, Alexandra Y., and Robert M. W. Dixon (eds.). 2006. *Serial Verb Constructions: A Cross-Linguistic Typology*. Oxford: Oxford University Press.

- Aronoff, Mark. 1994. *Morphology by Itself*. Cambridge, Mass: MIT Press.
- Ascoli, Graziadio I. 1896. "Un problema di sintassi comparata dialettale." *Archivio glottologico italiano* 14: 453–468.
- Ascoli, Graziadio I. 1901. "Appendice all'articolo Un problema di sintassi comparata dialettale." *Archivio glottologico italiano* 15: 221–225.
- Biberauer, Theresa, and Ian Roberts. 2012. "Towards a Parameter Hierarchy for Auxiliaries: Diachronic Considerations." Paper presented at DGFS. Cambridge: University of Cambridge.
- Cardinaletti, Anna, and Giuliana Giusti. 2001. "Semi-lexical Motion Verbs in Romance and Germanic." In *Semi-lexical categories*, ed. by Norbert Corver, and Henk Van Riemsdijk, 371–414. Berlin: De Gruyter. <https://doi.org/10.1515/9783110874006.371>
- Cardinaletti, Anna, and Giuliana Giusti. 2003. "Motion Verbs as Functional Heads." In *The Syntax of Italian Dialects*, ed. by Christina Tortora, 31–49. New York: Oxford University Press.
- Cardinaletti, Anna, and Giuliana Giusti. To appear. "Multiple Agreement in Southern Italian Dialects." In *Linguistic Variations: Structure and Interpretation – Studies in Honor of M. Rita Manzini*, ed. by Ludovico Franco, and Paolo Lorusso. Berlin: De Gruyter.
- Cardinaletti, Anna, and Ur Shlonsky. 2004. "Clitic positions and Restructuring in Italian." *Linguistic Inquiry* 35 (4): 519–557. <https://doi.org/10.1162/0024389042350523>
- Cinque, Guglielmo. 2006. *Restructuring and Functional Heads: The Cartography of Syntactic Structure*, 4. Oxford: Oxford University Press.
- Cruschina, Silvio. 2013. "Beyond the Stem and Inflectional Morphology: An Irregular Pattern at the Level of Periphrasis." In *The Boundaries of Pure Morphology*, ed. by Silvio Cruschina, Martin Maiden, and John Charles Smith, 262–283. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199678860.003.0014>
- Di Caro, Vincenzo N. 2015. "Syntactic Constructions with Motion Verbs in Some Sicilian Dialects: A Comparative Analysis." M.A. Thesis, Ca' Foscari University of Venice.
- Di Caro, Vincenzo N., and Giuliana Giusti. 2015. "A Protocol for the Inflected Construction in Sicilian Dialects." *Annali di Ca' Foscari, Serie Occidentale* 49: 393–422.
- Dressler, Wolfgang U., and Anna Thornton. 1991. "Doppie basi e binarismo nella morfologia italiana." *Rivista di Linguistica* 3 (1): 3–22.
- Ledgeway, Adam. 2015. "From Coordination to Subordination: the Grammaticalisation of Progressive and Andative Aspect in the Dialects of Salento." In *Coordination and Subordination*, ed. by Adriana Cardoso, Ana Maria Martins, Sandra Pereira, Clara Pinto, and Fernanda Pratas, 157–184. Oxford: Oxford University Press.
- Leone, Alfonso. 1973. "Vattel'a pesca, vieni a piglialo." *Lingua Nostra* 34: 11–13.
- Leone, Alfonso. 1978. "Sullo scadimento semantico di andare." *Lingua Nostra* 39: 50–54.
- Maiden, Martin. 2004. "When lexemes become allomorphs – on the genesis of suppletion." *Folia Linguistica* 38: 227–56. <https://doi.org/10.1515/flin.2004.38.3-4.227>
- Manzini, M. Rita, and Leonardo M. Savoia. 2005. *I dialetti Italiani e Romanci: Morfosintassi Generativa*. Alessandria: Edizioni dell'Orso.
- Rohlf, Gerhard. 1969. *Grammatica storica della lingua italiana e dei suoi dialetti, vol. 3: Sintassi e formazione delle parole*. Turin: Einaudi.
- Sornicola, Rosanna. 1976. "Vado a dire o vaiu a ddico problema sintattico o problema semantico?" *Lingua Nostra* 37 (3–4): 65–74.
- Sorrento, Luigi. 1950. *Sintassi romanza: Ricerche e prospettive*. Milan: Istituzione editoriale cisalpino.
- Stefanini, Ruggero. 1970. "Imperativo per infinito in fiorentino." *Lingua Nostra* 31: 19–20.

- Thornton, Anna. 2007. "Is There a Partition in the Present Indicative of Italian Irregular Verbs?" *Annali Online di Ferrara – Lettere* 2: 43–61.
- Zanuttini, Raffaella. 2008. "Encoding the Addressee in the syntax: Evidence from English imperative subjects." *Natural Language and Linguistic Theory* 26 (1):185–218.
<https://doi.org/10.1007/s11049-007-9029-6>

Ma non era rosso? (But wasn't it red?)

On counter-expectational questions in Italian

Alessandra Giorgi

Ca' Foscari University of Venice

In this article I investigate the properties of counter-expectational surprise yes-no questions in Italian, introduced by the adversative particle *ma*. These structures minimally contrast with surprise exclamations. I provide an analysis of the left periphery able to explain the observations concerning the distribution of *ma*, for instance that it must precede all other items in the clause and cannot be embedded. I propose that *ma* is a discourse head, projecting a syntactic structure analogous to that of normal syntactic heads. Discourse heads, however, connect separate sentences, which can also be uttered by different speakers, provided that they belong to the same context. I also add some brief remarks on the so-called *expletive* negation appearing in these cases.

Keywords: surprise questions, counter-expectational discourse, left periphery, expletive negation

1. Introduction

Surprise questions – or using a more appropriate terminology, introduced by Vicente (2010), *counter-expectational* questions – are especially interesting because they constitute a challenge for the very notion of sentence grammar. Their form and syntax are in fact determined to a large extent on the basis of extra-sentential properties. The aim of this work is to show how the extra-sentential issues can be handled in a Minimalist framework and combined successfully with a Cartographic model of sentence grammar, capturing the interactions between discourse and sentence in a principled way. The structure I propose for these constructions is a multi-sentential one, i.e. I'll argue that counter-expectational questions are *discourses*.

In this work, I consider the same set of data analyzed in Giorgi (2016b). In that paper, I focused mostly on their temporal interpretation; here, I extend the

analysis to the other syntactic properties, in particular word order phenomena and information structure. I compare this kind of questions with exclamation, because they minimally contrast for the relevant properties.

2. The data

Counter-expectational questions in Italian are introduced by the adversative particle *ma* (but). Note that for some speakers, the presence of *ma* is optional, as will be better discussed below in Section 5. Consider the following examples:

Scenario I

Mary calls me on the phone and tells me that she has a new red dress to wear at tonight's party. When I meet her at the party, I see that she has a blue dress. I'm surprised and say (from Giorgi, 2016b, Example 1):

- (1) *Ma non era rosso?*
 but not was-IMPF red
 'But wasn't it red?'

Scenario II

Mary informs me that she is going to buy her wedding dress. Later she shows me her purchase and I see that it is a red gown, an unusual color for this kind of dress. I may react by saying (from Giorgi 2016b, Example 2):

- (2) *Ma è rosso!*
 But is-IND red!
 'But it's red!'

The main properties that can be associated with (1) are the following:

- (3) a. There is a characteristic falling-raising intonation.
 b. The sentence is accompanied by gestures of the hands and of the head, plus brow raising.
 c. Presence of the imperfect form of the indicative mood.
 d. Presence of the particle *ma*.
 e. Presence of negation.

Intonation and gestures are usually considered properties pertaining to modules of grammar other than syntax (phonology, pragmatics). Moreover, surprise questions and exclamations have been analyzed so far as mono-sentential constructions.¹

In this paper, I challenge this view with respect to both issues. In previous work (Giorgi 2010), I already argued that the syntactic representation must be enriched by means of the representation in the C-layer of the speaker's temporal and spatial coordinates. I also argued (Giorgi 2014, 2016a) that certain properties concerning the prosodic contour of a sentence must be read off the syntax. In particular, I analyzed certain types of parentheticals, such as those introducing Free Indirect Discourse and Quotations, and proposed that their syntactic representation includes prosody-oriented heads, responsible for the characteristic *comma intonation* (cf. Selkirk (2005)) assigned to these structures. Giorgi (2015) also argues in favor of a similar theory for Italian Clitic Left Dislocation (henceforth, CLLD) contrasting it with Hanging Topics (henceforth, HT). In that paper, I proposed an analysis of HT as *discourses*, i.e. I proposed that what apparently looks like a single sentence, is actually composed of two different sentential units – one constituted by the hanging topic phrase, and the other by the main sentence – connected by means of a silent discourse head.²

Here I capitalize on this view and argue that the syntactic representation of counter-expectational yes-no questions must be enriched with information traditionally considered as belonging to other modules of grammar and, furthermore, that they must be represented as *bi-sentential structures*, i.e. *discourses*.

As far as the intonation is concerned, Lepschy (1978) already identified 5 tunes for Standard Italian: a falling tune characterizing statements, a rising tune for yes/no questions, a level tune for uncertainty, a rising-falling tune as a contradiction contour and a falling-rising tune expressing doubt or surprise, the last one being the relevant one in these cases.³

With respect to gestures, consider that their presence is obligatory, even if they can slightly vary across speakers. Namely, these sentences become strongly odd if pronounced with a neutral facial expression and with non-moving hands – as for

1. See Vicente (2010), Munaro and Obenauer (2002), Portner and Zanuttini (2003), Obenauer (2004), Delfitto and Fiorin (2014a, 2014b), a.o.

2. See Section 4.2 below for a discussion of these cases.

3. A discussion of the intonation of these constructions is outside the scope of this work. I just want to point out the relationship between the presence of a certain intonation and a peculiar syntax, strongly tied to each other.

instance with one's hands in the pockets. Adopting Schlenker's terminology, these gestures are co-speech ones, in that they accompany the whole sentence.⁴

3. The imperfect

In this section I briefly summarize the discussion of Giorgi (2016b) concerning the presence of the imperfect in counter-expectational questions, as opposed to exclamations, which exhibit a non-imperfect form of the indicative.

As a first observation, note that the imperfect is a well-behaved indicative form, in that in complement clauses it resists complementizer deletion, which is on the contrary a property of the subjunctive mood:⁵

- (4) *Gianni ha detto *(che) ieri alle tre mangiava un panino*
Gianni said that yesterday at three (he)-eat-IMPF a sandwich

In order to illustrate the contribution of the imperfect to the constructions at issue, I will show what happens if the distribution of the imperfect vs. the non-imperfect indicative in (1) and (2) is reversed. Consider the following case (from Giorgi (2016b), Example 38):

- (5) *Ma non è rosso?*
But isn't-IND it red?

In Scenario I, a sentence with a present indicative like (5) is infelicitous. It is appropriate in Scenario III:

Scenario III

Mary, pointing to a dress exposed in a window, tells Paul: "How beautiful that blue dress!" and Paul might answer: "But isn't it red?", because he is seeing it as red and not as blue.

In this case, the speaker assumes that the dress is red and asks for an explanation.

4. Schlenker (2015, abstract), "[...] We argue that some co-speech gestures should be analyzed within a presuppositional framework, but with a twist: an expression *p* co-occurring with a co-speech gesture *G* with content *g* comes with the requirement that the local context of *p* should guarantee that *p* entails *g*; we call such assertion-dependent presuppositions 'cosuppositions' [...]."

5. With some exceptions, which I do not discuss here, because they are not relevant for the purposes of this work. For a full discussion, see Giorgi (2010).

Conversely, in Scenario II an exclamative sentence with an imperfect like (6) is infelicitous (from Giorgi (2016b), Example 39):

- (6) *Ma era rosso!*
But it was-*IMPF* red!

It is appropriate in Scenario IV:

Scenario IV

The speaker knows that I will wear a certain dress, which he remembers as red. When he sees me, he recognizes the dress as the one he remembers, but the dress is presently of a different color.

The presence of the imperfect here expresses a temporal value, in that the speaker says that the dress was red at a previous time, and then changed, for instance by dying.

Note also that there is an important semantic difference between the question with the imperfect and the one with the indicative. The sentence with the imperfect, repeated here, is a *special* question, according to the terminology by Obenauer (2004):

- (7) *Ma non era rosso?* (=1)
But not was-*IMPF* red
'But wasn't it red?'
- (8) *#Sì, era rosso*
Yes, it was red
- (9) *#No, no era rosso*
No, it wasn't red

A yes-no answer would not make any sense at all: the speaker wants an explanation with respect to the dress color, because it does not meet her expectations. On the contrary, the question with the indicative, should indeed be answered in the 'normal' way:

- (10) *Ma non è rosso?* (=5)
But isn't-*IND* it red?
- (11) *Sì, è rosso (mi sono sbagliata)*
Yes. It's red (I was wrong)
- (12) *No, non è rosso (è la luce)*
No, it isn't red (it is the light)

In what follows I provide an explanation for these observations.⁶

Recall first that the imperfect is an anaphoric verbal form, i.e. there is no direct anchoring to the utterance time. Consider the following examples:⁷

- (13) *#Luca faceva i compiti.*
Luca do-IMPF homework.
- (14) *Ieri alle tre Luca faceva i compiti.*
Yesterday at three Luca do-IMPF homework
'Yesterday at three Luca was doing homework'
- (15) *Cosa faceva Luca alle tre?*
What was Luca doing at three?
- (16) *Luca faceva i compiti*
Luca do-IMPF homework
'Luca was doing homework'

In these examples, the relationship of the imperfect with the utterance time must be mediated by the presence of a temporal reference, either present in the same sentence, or in the preceding discourse/context. Giorgi and Pianesi (2001) argued that exactly this property determines the distribution of this verbal form in *fictional* and *dream* contexts, where it is obligatory. In what follows, I'll give a brief overview of various phenomena.

Consider first the so-called *imperfâit prelude*, typically used by children while planning a new game (cf. Vet 1983):

- (17) *Facciamo che io ero il re e tu la regina*
Let's pretend that I am-IMPF the king and you the queen

Analogously, consider stage instructions:

- (18) *A questo punto il ladro usciva e tu lo inseguivi.*
At this point the thief leave-IMPF and you follow-IMPF him

6. The literature on the imperfect is very rich and I'll provide here only a brief discussion limited to purposes of this work, in the framework of Giorgi and Pianesi (1997, 2001). See also, for different perspectives on the issue, Delfitto and Bertinetto (1995), Delfitto (2004) and Ippolito (2000).

7. Note that in these cases the imperfect has a past temporal value. However, I will illustrate in the text other usages in which the imperfect does not have a temporal value at all.

In both cases the imperfect is not referring to a past event and is obligatory. Consider for instance sentences (17) and (18) with a present perfect, instead of an imperfect:⁸

- (19) **Facciamo che io sono stato/fui il re e tu la regina.*
 Let's pretend that I was-PRES PERF the king and you the queen.
- (20) **A questo punto il ladro è uscito e tu lo hai seguito.*
 At this point the thief left-PRES PERF and you followed-PRES PERF
 him.

These sentences cannot be taken to be stage instructions, but descriptions of past events.

The imperfect is also the form used in narrative contexts, story-telling and fiction:

- (21) *Il ladro passeggiava nervosamente. Qualcosa era andato storto...*
 The thief walked-IMPF nervously. Something had-IMPF gone wrong...
 (from Giorgi & Pianesi (2001), Example 53)

Again, there is no way in which these events can be interpreted as past with respect to the utterance time. Moreover, the imperfect is the verbal form embedded under the verb *sognare* (dream):

- (22) *Luca ha sognato che Paolo vinceva la gara.*
 Luca dreamed that Paolo won-IMPF the race.

The winning of the race is *the content of the dream* and it is not located in the past present or future with respect to the dream itself. I.e., it is not the case that Luca dreamed of an event located in his past, contrasting with the following example:

- (23) *Luca ha detto che Paolo ha vinto la gara.*
 Luca said that Paolo won-PRES PERF the race

8. In Central and Northern Italy, the present perfect is the form normally used to express pastness. In Southern Italy, a simple past would be used. In the examples in the text, I adopt my own variety (Central Italy).

In Example (23), the winning of the race is temporally located in the past with respect to the utterance time.⁹

In all the cases illustrated above, the imperfect event is not anchored to the utterance time. It contrasts with the other forms of the indicative, for instance the present perfect, which on contrary must express a temporal value of past-ness with respect to the utterance time.

These observations account for the obligatory presence of the imperfect in counter-expectational questions, in that in these cases the embedded verbal form cannot be anchored to the indexical context: the redness of the dress is not a fact, but only the speaker's expectation and for this reason, the predicate is not anchored to the context where the speaker is located.

Conversely, a non-imperfect indicative requires anchoring to the utterance time. Therefore, in this case, for the sentence to be felicitous, the dress must indeed be red. In the case of exclamative sentences the opposite holds. The speaker's surprise is due to the fact that, according to her, the dress is red. Therefore, an indexical verbal form is required, as for instance a present perfect.

9. The imperfect does not exhibit the Double Access Reading (DAR). For reasons of space, it is impossible to provide a discussion of this issue here, and refer the reader to the previous literature. However, it might be worth mentioning, because it is an interesting argument in favor of the analysis provided in the text. The basic example for the DAR includes a present tense embedded under a past verbal form:

- i. *Gianni ha detto che Maria è incinta*
Gianni said that Maria is pregnant

For this sentence to be felicitous, the state of pregnancy must hold both at the time of the saying and at utterance time, so that the following sentence is infelicitous:

- ii. *#Due anni fa, Gianni ha detto che Maria è incinta*
Two years ago, Gianni said that Maria is pregnant

However, there is no DAR with an embedded imperfect:

- iii. *Gianni ha detto che Maria era incinta*
Gianni said that Maria was(IMPF) pregnant

In this case the state of pregnancy is simultaneous with the saying and does not have to hold at time of the utterance, to the extent that the following sentence contrasts with (ii) above:

- iv. *Due anni fa, Gianni ha detto che Maria era incinta*
Two years ago, Gianni said that Maria was(IMPF) pregnant

This contrast has been explained by Giorgi and Pianesi (1997) and Giorgi (2010) on the basis of the idea that the imperfect is not an indexical verbal form, which is the point I am stressing in the text.

4. The distribution and properties of the particle *ma* in counter-expectational contexts

4.1 Generalities

Ma cannot appear in embedded contexts, independently of its position with respect to the complementizer *che*:

(24) **Gianni ha detto che ma non era rosso*
Gianni said that but it wasn't-IMPF red

(25) **Gianni ha detto che ma è rosso*
Gianni said that but it is-IND red

Ma is not itself a complementizer, as shown by the ungrammaticality of the following examples:

(26) **Gianni ha detto ma non era rosso*
Gianni said but it wasn't-IMPF red

(27) **Gianni ha detto ma è rosso*
Gianni said but it is-IN red

In non-counter-expectational contexts, *ma* appears in a conjunction:

(28) *Maria ha trenta anni, ma ne dimostra venti*
Maria is thirty, but she looks twenty

(29) *Maria è stanca, ma felice*
Maria is tired, but happy

Note that (28) and (29) can also be realized as discourses, i.e. the two parts could be uttered by different speakers:

(30) A: *Maria ha trenta anni*
Maria is thirty

(31) B: *(Sì), ma ne dimostra venti*
(Yes), but she looks twenty

(32) A: *Maria è stanca*
Maria is tired

(33) B: *(Sì), ma (è) felice*
(Yes), but (she is) happy

In these cases, the presence of the first conjunct, either in the same sentence, or in a discourse is obligatory. A non-counter-expectational constructions, without the first conjunction is infelicitous:

(34) *#Ma ne dimostra venti*
But she looks twenty

(35) *#Ma è felice*
But she is happy

4.2 The co-occurrence of *ma* with clitic left dislocation, focus and hanging topic

In this section I consider the distribution of *ma* with respect to the other items of the left periphery. Roughly speaking, we can say that in Italian the left periphery is the syntactic portion of the clause on the left of the subject. The particle *ma* indeed occurs in this area, hence, in principle it should qualify as component of the left periphery.

According to Rizzi (1997), the left periphery of the clause is structured as follows:¹⁰

(36) C-FORCE > TOP* > FOCUS > TOP* > FIN

In Italian the complementizer *che* (that), introducing finite clauses, appears in C-FORCE, whereas the complementizer introducing non-finite clauses appears in FIN. TOP is usually taken to be a (clitic) left dislocated phrase and FOCUS is a contrastive focus. The asterisk on the right of TOP means that it is possible to recur on the topic position. Therefore, according to this hypothesis, a clitic left dislocated phrase can occur anywhere in the left periphery, both on the left and on the right of Focus, provided that it occurs on the left of the complementizer. Focus, on the contrary is a unique position. Concluding, in principle, if *ma* is a head of the left periphery, even if of a type not yet identified, we should be able to ascertain its position by looking at its distribution with respect to Focus and Topic.

A clitic left dislocated phrase cannot appear on the left of *ma*, both in the case of questions and in the case of exclamatives:¹¹

10. Many scholars addressed these issues in the framework of the so-called *cartographic approach*. See among the others Belletti (2004) and Rizzi (2004). I refer the reader to the relevant literature.

11. For some speakers, Example (37) is strongly ungrammatical, while for other ones, less so. According to my intuition, it is quite bad: '?*'. The reasons for these differences among speakers are not clear yet to me.

- (37) ??*A Luca, ma non gli avevi dato un libro?*
To Luca, but not (you) to him(CL) had(IMPf) given a book?
- (38) **A Luca ma gli hai dato un libro!*
To Luca but (you) to him have given a book!

When on the right of *ma*, a CLLD item is grammatical, as shown by the following examples:

- (39) *Ma a Luca, non gli avevi dato un libro?*
But to Luca, not to him(CL) (you) had(IMPf) given a book?
- (40) *Ma a Luca, gli hai dato un libro!*
But to Luca (you) to him have given ice-cream!

Let's compare now CLLD with contrastive focus (in capital letters). The focused phrase cannot precede the particle *ma*. Consider the following examples:

- (41) **UN LIBRO (non un vestito) ma non avevi comprato a Maria?*
A book (not a dress) but (you) not had bought to Maria?
- (42) **UN LIBRO ma hai comprato! (non un vestito)*
A book but you bought! (not a dress)

Finally, a focused phrase on the right of *ma* is ungrammatical with the interrogative structure:

- (43) **Ma UN LIBRO (non un vestito) non avevi comprato a Maria?*
But a book (not a dress) (you) not had bought to Maria?

On the contrary, in exclamatives a contrastive focus following *ma* is grammatical:

- (44) *Ma UN LIBRO hai comprato! (non un vestito)*
But the A BOOK you bought! (not a dress)
'But you bought a book! (not a dress)'

Hence, neither CLLD nor Focus can precede *ma*. With the exception of Example (43), they can follow it. With respect to (43), consider that this judgment is expected, given that in Italian it is in general impossible to have questions with a co-occurring contrastive focus in the left periphery:

- (45) **IL LIBRO (non i biscotti) hai dato a Maria?*
The book-foc (not the cookies) (you) have given to Maria?

Note that in the literature on these issues, it has been long argued that contrastive focus is a moved phrase, giving rise to an operator-variable structure, whereas CLLD is base generated.¹²

Therefore, we can conclude that the position on the left of *ma* is available neither to moved phrases, nor to base generated ones, whereas the one on its right is in general accessible.

Consider now a Hanging Topic phrase. Hanging Topic is grammatical on the left of *ma* in both constructions:¹³

(46) *Mario, ma non gli avevi comprato un libro?*
 Mario, but (you) non to him-cl had-IMPF bought a book?

(47) *Mario, ma gli hai comprato un libro!*
 Mario, but (you) to him-cl bought an a book!

In the next section I provide an explanation for these observations.

5. Sentence and discourse

Giorgi (2015) points out that there are many differences between CLLD and HT, which cannot be accounted for by simply hypothesizing that they are ‘variants’ of the same constructions. Capitalizing on an observation by Cinque (2008), Giorgi (2015) proposes that HT and the sentence following it, though apparently constituting a single sentence, are actually a discourse, whereas this is not the case for CLLD. Consider the following HT example:¹⁴

(48) *Paolo, il professore gli ha dato un bellissimo voto*
 Paolo, the professor to him-cl has given a very good mark

Giorgi (2015) proposes that there is a silent discourse head *DIS*, connecting the two parts of the discourse. HT appears in the specifier position, whereas the sentence in the complement one. Therefore, the structure is the following:

12. See, among the others, Rizzi (1997), Cinque (1990), Frascarelli (2000).

13. In this case HT can be distinguished from CLLD because the phrase appearing on the left though corresponding to a dative argument – cf. the dative clitic *gli* (to him) appearing in the clause – is not accompanied by the dative preposition *a* (to), as in *a Paolo* (to Paolo). This is the most evident difference between HT and CLLD. Again, a full discussion of this issue is not possible in this work.

14. For reasons of space, I cannot reproduce here the relevant discussion in Giorgi (2015), where it is shown that HT, in spite of its name, does not behave at all as an ordinary topic.

- (49) [_{DISCOURSE} Paolo [_{DIS} [*il professore gli ha dato un bellissimo voto*]]
 Paolo the professor gave him-cl a very good mark

In this work, I extend this analysis to *ma* constructions and show that it can capture the data discussed above. In particular, I claim that *ma* is the head of a discourse taking the interrogative clause as its complement, whereas the specifier position is empty:

- (50) [_{DISCOURSE} ... [_{DIS} *ma* [*non era rosso?*]]
 but wasn't it red?

- (51) [_{DISCOURSE} ... [_{DIS} *ma* [*è rosso!*]]
 but is red!

The dots under the specifier stand for the silent portion of the construction, namely, the expectation for the dress to be, or to be not, red.¹⁵

Note that in this way it is possible to account for the observations in (28)–(35). Consider for instance example (28), or (29), according to this proposal, the structure is the following:

- (52) [_{DISCOURSE} *Maria ha trenta anni* [_{DIS} *ma* [*ne dimostra venti*]]
 Maria is thirty but she looks twenty

Where specifier and complement¹⁶ can also be uttered by different speakers. The presence of the appropriate intonation and gestures is required when the specifier is missing, i.e. in counter-expectational constructions. As mentioned above in Section 1, for some speakers in these cases the discourse head *ma* can be silent, provided that the counter-expectational value can be retrieved by means of a very strong intonation and emphatic gestures. Consider also that, as noted in Examples (30)–(33), when the two conjuncts are uttered by different speakers, the affirmative particle *sì* (yes) can appear. According to this proposal, *sì* (yes) occupies the specifier position of the DIS projection:

15. A reviewer raises the very important question concerning the notion of *sentence* adopted in this work. Certainly, it cannot be a simple notion of “syntactic sentence” in the traditional sense – notion already challenged in the last twenty years in the studies on parentheticals – but must distinguish among several notions: an intonational one, a pragmatic one and a syntactic one, depending on the object of the analysis. Unfortunately, for reasons of space, this issue cannot be addressed here.

16. There is no principled requirement in fact for the specifier and complement of a discourse head to share the same speaker. All is required in these cases is that they share the same *context*, i.e. simplifying, the occasion in which a speech act takes place, which permits the identification of the participants in the conversation.

- (53) [_{DISCOURSE} *Si* [_{DIS} *ma* [*ne dimostra venti*]]
 Yes but she looks twenty

Si (yes) in this case refers anaphorically to the first conjunct uttered by the other speaker.

Let's go back now to the distribution of CLLD and Focus. The unavailability of both moved and unmoved phrases on the left of *ma*, is explained by means of the consideration that this area does not belong to the sentence. *Ma* is a discourse head and defines the boundary between sentences, hence all the phrases belonging to that sentence must appear on its right.

Hanging Topic, on the contrary, as illustrated above, gives rise to a discourse as well. Hence, for sentence (46) and (47) the structure is the following:

- (54) [_{DISCOURSE} *Mario* [_{DIS} \emptyset [_{DISCOURSE} ... [_{DIS} *ma* [*non gli avevi comprato un libro?*]]]]]
 Mario but didn't you buy a book to
 him?

- (55) [_{DISCOURSE} *Mario* [_{DIS} \emptyset [_{DISCOURSE} ... [_{DIS} *ma* [*gli hai comprato un libro!*]]]]]
 Mario but you bought a book to
 him!

Here we have a complex discourse with a high empty discourse head. The HT appears in its specifier position and the *ma* construction is its complement. On its turn, *ma* projects a discourse phrase, with an empty specifier position, as discussed above.

I conclude this section with a few words on the presence of negation. Delfitto and Fiorin (2014b), point out that negation in surprise exclamatives is expletive, i.e. it does not contribute its canonical meaning as a propositional operator, and only prompts a positive reply. *Prima facie* the same analysis could be proposed for the counter-expectational constructions, even if the authors do not consider exactly this type. My proposal is that at least in these cases however, negation is not expletive, but a real 'normal' sentential negation.

In Section 1, I listed the properties of these constructions, let me summarize here what has been observed so far: Intonation and gestures, together with *ma*, express the counter-expectational nature of the structure and request for more information; the imperfect stresses the fact that the predicate holds with respect to the speaker's expectations, and not in the real world. Hence, these components roughly speaking express the following meaning: "X contradicts my expectations, I require further information". Once these components – intonation, gestures, interrogative form, *ma*, and imperfect – are stripped off the sentence, what we are

left with *is dress not red*, which is indeed true. From this perspective, the negation is not expletive at all.¹⁷

6. Conclusions

In this article, I proposed an enlarged view of grammar, able to include certain discourse phenomena as well. I proposed that *ma* is a discourse head, which projects a syntactic structure analogous to that of normal syntactic heads appearing in the sentence. Discourse heads connect separate sentences, which can also be uttered by different speakers, provided that they belong to the same context. Further study is indeed required in order to consider possible developments of this view to other discourse heads and other constructions.

References

- Belletti, Adriana (ed.). 2004. *Structures and Beyond. The Cartography of Syntactic Structures*, 3. Oxford University Press: New York.
- Cinque, Guglielmo. 1990. *Types of A-bar Dependencies*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 2008. "Two Types of Nonrestrictive Relative Clauses." In *Empirical Issues in Syntax and Semantics 7*, ed. by Olivier Bonami, and Patricia Cabredo Hofherr, 99–137. Paris: CSSP.
- Delfitto, Denis. 2004. "On the logical form of imperfective aspect." In *The syntax of time*, ed. by Jacqueline Guéron, and Alexander Lecarme, 115–142. Cambridge, MA: MIT Press.
- Delfitto, Denis, and Pier Marco Bertinetto. 1995. "A case study in the interaction of aspect and actionality: The imperfect in Italian." *Temporal Reference, Aspect and Actionality, vol. I, Semantic and Syntactic Perspectives*, ed. by Pier Marco Bertinetto, Valentina Bianchi, James Higginbotham, and Mario Squartini, 125–142. Turin: Rosenberg & Sellier.
- Delfitto, Denis, and Gaetano Fiorin. 2014a. "Exclamatives: Issues of syntax, logical form and interpretation." *Lingua* 152, 1–20. <https://doi.org/10.1016/j.lingua.2014.09.003>
- Delfitto, Denis, and Gaetano Fiorin. 2014b. "Negation in Exclamatives." *Studia Linguistica* 68 (3): 284–327. <https://doi.org/10.1111/stul.12024>
- Frascarelli, Mara. 2000. *The Syntax-Phonology Interface in Focus and Topic Constructions in Italian*. Dordrecht: Kluwer. <https://doi.org/10.1007/978-94-015-9500-1>
- Giorgi, Alessandra. 2010. *About the Speaker: Towards a Syntax of Indexicality*. Oxford: Oxford University Press.
- Giorgi, Alessandra. 2014. "Prosodic Signals as Syntactic Formatives in the Left Periphery." In *On Peripheries: Exploring Clause Initial and Clause Final Positions*, ed. by Anna Cardinaletti, Guglielmo Cinque, and Yoshio Endo, 161–188. Tokyo: Hituzi Syobo.

17. The possibility of generalizing this conclusion to other cases is a topic for further research.

- Giorgi, Alessandra. 2015. "Discourse and the syntax of the left periphery: clitic left dislocation and hanging topic." In *Discourse-oriented syntax*, ed. by Josef Bayer, Roland Hinterhölzl, and Andreas Trotzke, 229–250. Amsterdam: John Benjamins.
<https://doi.org/10.1075/la.226.10gio>
- Giorgi, Alessandra. 2016a. "Integrated Parentheticals in Quotations and Free Indirect Discourse." In *Indirect Reports and Pragmatics*, ed. by Alessandro Capone, Ferenc Kiefer, and Franco Lo Piparo, 471–488. New York: Springer. https://doi.org/10.1007/978-3-319-21395-8_21
- Giorgi, Alessandra. 2016b. "On the temporal interpretation of certain surprise questions." *SpringerPlus* 5 (1): 1390. <https://doi.org/10.1186/s40064-016-2951-5>.
- Giorgi, Alessandra, and Fabio Pianesi. 1997. *Tense and Aspect: From semantics to Morphosyntax*. New York: Oxford University Press.
- Giorgi, Alessandra, and Fabio Pianesi. 2001. "Imperfect Dreams. The temporal dependencies of fictional predicates." *Probus* 13: 31–68. <https://doi.org/10.1515/prbs.13.1.31>
- Ippolito, Michela. 2000. "Imperfect modality." Paper presented at the International Round Table *The Syntax of Tense and Aspect*, Université de Paris VII, France.
- Lepschy, Giulio. 1978. *The Italian Language Today*. London: Hutchinson.
- Munaro, Nicola, and Hans-Georg Obenauer. 2002. "On the semantic widening of underspecified wh-elements." In *Current Issues in Generative Grammar*, ed. by Manuel Leonetti, Olga Fernández Soriano, and Victoria Escandell Vidal, 165–194. Madrid: Universidad de Alcalá de Henares / UNED / Universidad Autónoma de Madrid.
- Obenauer, Hans-Georg. 2004. "Nonstandard wh-questions and alternative checkers in Pagotto." In *Syntax and Semantics of the Left Periphery*, ed. by Horst Lohnstein, and Susanne Trissler, 343–384. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110912111.343>
- Portner, Paul, and Raffaella Zanuttini. 2003. "Exclamative Clauses: At the Syntax-Semantic Interface." *Language* 79 (1): 39–81. <https://doi.org/10.1353/lan.2003.0105>
- Rizzi, Luigi. 1997. "The fine structure of the left periphery." In *Elements of Grammar*, ed. by Liliane Haegeman, 281–337. Dordrecht: Kluwer.
https://doi.org/10.1007/978-94-011-5420-8_7
- Rizzi, Luigi (ed.). 2004. *The Structure of IP and CP. The Cartography of Syntactic Structures, 2*. New York: Oxford University Press.
- Schlenker, Philippe. 2015. "Gesture Projection and Cosuppositions." Ms., Institut Jean-Nicod and New York University.
- Selkirk, Lisa. 2005. "Comments on intonational phrasing." In *Prosodies: With Special Reference to Iberian*, ed. by Sonia Frota, Marina Cláudia Vigário, and Maria João Freitas, 11–58. Berlin, Mouton de Gruyter.
- Vet, Co. 1983. From tense to modality, In *Studies in Modeltheoretic Semantics*, ed. by Alice G. B. Ter Meulen, 192–207. Dordrecht: Foris.
- Vicente, Luis. 2010. "On the syntax of adversative coordination." *Natural Language and Linguistic Theory* 28 (2): 381–415. <https://doi.org/10.1007/s11049-010-9094-0>

Dime una cosa: Are wh-in-situ questions different in Spanish?

Evidence from intonation

Carolina González and Lara Reglero

Florida State University

This project provides a preliminary examination of the intonational characteristics of information-seeking, echo-repetition and echo-surprise wh-in-situ questions in Spanish. The analysis of 120 questions from 4 female participants from Northern Peninsular Spanish shows that information-seeking questions have a relatively reduced tonal range, and they are the most likely to have final rising contours. On the other hand, echo-repetition and echo-surprise questions differ mostly in that the former have lower local tonal ranges than the latter. These findings provide support for previous syntactic analyses that argue for differences between information-seeking and echo-questions on one side, and echo-repetition and echo-surprise questions on the other.

Keywords: Peninsular Spanish, wh-in-situ, information-seeking questions, echo questions, intonation, marked stress, tonal range, final contours, duration

1. Introduction

As is well-known, languages display different patterns when it comes to wh-question formation. In wh-fronting languages, as in English (1), the wh-word undergoes overt movement to the Spec of CP from its base-generated position (Chomsky 1957, 2000, Pesetsky 1982, May 1985, Lasnik and Saito 1992):¹

(1) *What did John buy?*

1. See Ginzburg and Sag (2001), Pires and Taylor (2007), and Reglero (2007) for the availability of the wh-in-situ strategy in English with the true question, non-echo reading.

In wh-in-situ languages, as in Chinese, the wh-word is typically pronounced in the same position as its non-wh counterpart (2) (Huang 1982, Aoun and Li 1993, Cole and Hermon 1998):

- (2) *John mai-le shenme?*
 John buy-ASP what
 ‘What did John buy?’

Finally, some languages such as French exhibit a ‘mixed’ system, allowing both wh-fronting (3a) and wh-in-situ (3b) (Chang 1997, Bošković 2000, Cheng and Rooryck 2000, Vergnaud and Zubizarreta 2002, Mathieu 2004, Déprez et al. 2013, a.o.).

- (3) a. *Qu’a acheté John?*
 what.has bought John
 ‘What has John bought?’
 b. *John a acheté quoi?*
 John has bought what
 ‘What has John bought?’

Spanish also exhibits a ‘mixed’ pattern for wh-question formation (4) (Jiménez 1997, Uribe-Etxebarria 2002, Etxepare and Uribe-Etxebarria 2005, Reglero 2007, Reglero and Ticio 2013). As in French, (4b) can be interpreted as a true question seeking new information (henceforth, ‘information-seeking’ wh-in-situ question). In addition, it can also be interpreted as an echo question in two well-defined contexts: (1) to request repetition of information (henceforth, ‘echo-repetition’), or (2) to convey surprise (henceforth, ‘echo-surprise’) (Reglero and Ticio 2013).^{2,3}

2. An anonymous reviewer points out that the echo reading can also be conveyed by wh-fronting in some languages. For example, in Pagotto, bare wh-phrases need to appear sentence-initially when they have a surprise reading (Munaro and Obenauer 1999, 2002; Obenauer 2004, 2006). Contreras (1999) also notes that echo wh-phrases can appear fronted in Spanish, as in (i) and (ii). Note that in these examples, and throughout this paper, capital letters indicate wh-words with an echo reading.

- (i) *¿A QUIÉN le diste ese libro?*
 to who CL gave that book
 ‘TO WHO did you give that book?’
 (ii) *¿Cómo que CUÁNDO llegué?*
 how that when arrived
 ‘I arrived WHEN?’

3. Sobin (2010) distinguishes between syntactic echo questions (typically in-situ), vs. pseudo echo questions (with wh-fronting and echo question intonation). In this paper, we focus on syntactic echo questions.

- (4) a. *¿Qué compró Juan?*
 what bought Juan
 ‘What did Juan buy?’
 b. *¿Juan compró qué?*

Although information-seeking wh-in-situ questions in Spanish are not as frequent as in French, they are well attested in some dialects, such as Northern Peninsular Spanish (Uribe-Etxebarria 2002, Etxepare and Uribe-Etxebarria 2005, a.o.). Various analyses have been proposed to explain the syntactic differences of information-seeking and echo questions in situ (Uribe-Etxebarria 2002, Etxepare and Uribe-Etxebarria 2005, Reglero 2007, Reglero and Ticio 2013, Chernova 2013). Several authors have pointed out intonational differences between the two, but there is no thorough acoustic description of these differences or how they interact with the syntactic component. Thus, this paper aims to provide a preliminary investigation on the intonation of wh-in-situ questions in Spanish, focusing on their main pragmatic readings: (1) information-seeking, (2) echo-repetition, and (3) echo-surprise.

The paper is organized as follows. Section 2 outlines the pragmatics and syntax of wh-in-situ questions in Spanish and reviews prior reports on their intonational characteristics. Section 3 describes the experimental design of our study, and Section 4 presents the results obtained. Finally, discussion and concluding remarks are offered in Section 5.

2. Wh-in-situ in Spanish: Pragmatics, syntax and intonation

This section discusses the pragmatic readings of Spanish wh-in-situ questions (2.1) and the syntactic analyses that have been proposed for them (2.2). It also reviews prior reports on their intonational differences (2.3).

2.1 Pragmatic readings

Information-seeking wh-in-situ questions request new information and can be used in out-of-the-blue contexts (Reglero 2004, 2007, Reglero and Ticio 2013).⁴ In contrast, echo questions are more heavily dependent on a previous context, and cannot be used out-of-the-blue. Reglero and Ticio (2013) argue that this difference

4. Cf. Jiménez (1997) and Uribe-Etxebarria (2002), who argue for a required context with a pre-established set of answers.

explains why the former can be introduced by *dime una cosa* ‘tell me something’ (5) while the latter cannot (6).⁵

- (5) *Dime una cosa: ¿María se tomó un té con quién?*
 tell.me one thing María CL had a tea with who
 ‘Tell me something: María had a tea with who?’
- (6) Speaker 1: *María se tomó un té con Cleopatra.*
 María CL had a tea with Cleopatra
 Speaker 2: **Dime una cosa: ¿María se tomó un té con QUIÉN?*
 tell.me one thing María CL had a tea with WHO

Furthermore, echo questions are ambiguous in that they can request repetition of information or convey surprise. In order to distinguish between the echo-repetition and echo-surprise readings, we follow Reglero and Ticio (2013) and propose the sequences *perdona, no te he oído bien* ‘excuse me, I didn’t hear you’ and *no me lo puedo creer* ‘I can’t believe it’, to disambiguate between the echo-repetition and echo-surprise readings, respectively. The dialogues in (7, 8) illustrate both interpretations.

- (7) Speaker 1: *María se tomó un té con Cleopatra.*
 María CL had a tea with Cleopatra
 Speaker 2: *Perdona, no te he oído bien: ¿María se tomó un té*
 excuse.me NEG CL have heard well María CL had a tea
 con QUIÉN?
 with WHO
- (8) Speaker 1: *María se tomó un té con Cleopatra.*
 María CL had a tea with Cleopatra
 Speaker 2: *No me lo puedo creer: ¿María se tomó un té con QUIÉN?*
 NEG CL CL can believe María CL had a tea with WHO

2.2 Syntactic analyses

There is a debate in the literature as to whether information-seeking wh-in-situ questions in Spanish involve overt movement or not.⁶ Uribe-Etxebarria (2002) and Etxepare and Uribe-Etxebarria (2005) propose that these questions have complex syntax involving two movement operations. Specifically, (i) the wh-word moves overtly to Spec CP, and (ii) the remnant IP moves to a topic-like position (XP).

5. Reglero and Ticio (2013) attribute the *dime una cosa* test to Ignacio Bosque (p.c.).

6. For movement and non-movement analyses for French, see Munaro et al. (2001) and Bošković (2000), respectively.

These two overt movement operations ensure that the in situ wh-word appears sentence-finally. A two-step derivation for example (4b) is given in (9) below.

- (9) a. [_{IP} *Juan compró qué*]
 Juan bought what
 b. [_{CP} *qué*_i [_{IP} *Juan compró t_i*]]
 c. [_{XP} [_{IP} *Juan compró t_i*]_j [_{CP} *qué t_j*]]

In contrast, Reglero (2007) argues for a non-movement approach in which the sentence-final position of the wh-word is derived as the result of its syntactic and phonological properties. Following insights by Zubizarreta (1998) and Stjepanović (1999, 2003), Reglero (2007) proposes that Spanish in situ wh-words must appear sentence-finally to receive main stress via the Nuclear Stress Rule. A derivation for (4b) is provided in (10). Note that the in situ wh-word *qué* ‘what’ is intrinsically marked [+F] because it corresponds to new information focus. The subject and the verb are [-F] because they are presupposed. As indicated in the derivation, the highest copies of *Juan* and *compró* ‘bought’ are pronounced. In contrast, the lowest copy of the [+F] *qué* ‘what’ gets pronounced to receive prosodic prominence in final position.⁷

- (10) [_{AgrSP} *Juan compró* [_{AgrOP} *qué compró* [_{VP} *Juan compró qué*]]]
 [-F] [-F] [+F] [-F] [-F] [-F] [+F]

More recently, Chernova (2013) and Reglero and Ticio (2013) have claimed that information-seeking and echo questions are completely different phenomena not only in terms of their pragmatics and semantics, but of their syntax as well. These authors note that echo questions exhibit a “Comp freezing” effect and do not need to appear in absolute sentence-final position, unlike information-seeking wh-in-situ questions.⁸ Following Sobin (2010), they argue that these questions have different complementizers with distinct feature compositions: C_{WH} and C_{EQ} (EQ

7. See Reglero and Ticio (2013) for an updated analysis adopting Kahnemuyipour’s (2009) system.

8. See Chernova (2013) and Reglero and Ticio (2013) for further examples and a more in-depth discussion.

stands for echo question), respectively.⁹ To illustrate, example (4b) would have the following structure (adapted from Chernova 2013):¹⁰

- (11) [_{CPEQ} [_{IP} *Juan compró QUÉ*]]
 C_{EQ} Juan bought what

2.3 Intonational properties

Although there are no prior acoustic studies on the intonation of *wh*-in-situ questions in Spanish, impressionistic reports suggest that information-seeking *wh*-in-situ questions have falling intonation and extra or ‘marked’ stress on the *wh*-word (Uribe-Etxebarria 2002, Reglero and Ticio 2013). Echo questions – particularly with a surprise reading – reportedly have rising or strong/sharp intonation, with ‘marked’ stress on the *wh*-word (Pope 1976, Contreras 1999, Sobin 2010, Chernova 2013). It is not clear in these sources what ‘extra’ or ‘marked’ stress is, and whether these terms refer to the same acoustic correlate or correlates.

In Slavic languages, and possibly cross-linguistically, there are impressionistic reports of stronger stress on the *wh*-word in echo-surprise contexts than in echo-repetition ones (Bošković 2002). For Spanish, Escandell (1999) claims that there is a circumflex pattern (rising-falling) for questions conveying surprise, versus a falling-rising intonation for those with a repetition reading.

3. Experimental design

This paper reports preliminary results from an acoustic study investigating the intonation of *wh*-in-situ questions in Northern Peninsular Spanish (NPS), where they are well attested (see for example Uribe-Etxebarria 2002). Based on the properties reviewed in Section 2, this study examines two main hypotheses. The first one is that information-seeking *wh*-in-situ questions will differ intonationally

9. To our knowledge, no one has proposed separate syntactic analyses for echo-repetition and echo-surprise questions in Spanish. However, Bošković (2002) found differences in grammaticality judgments in Slavic languages, with echo-surprise readings more acceptable in situ. He attributed this difference to the more D-linked nature of the echo-surprise *wh*-word (its value is known to the speaker and hearer). If this line of reasoning is correct, perhaps C_{EQ} could be further classified into C_{EQ} (repetition) and C_{EQ-D-Linked} (surprise) (see Cable (2010) for a similar distinction between non-D-linked and D-linked *wh*-phrases in English). For reasons of space, a more thorough analysis of this phenomenon is left for future research.

10. Irrelevant details are omitted from the derivation. See Chernova (2013) for a complete analysis.

from echo questions. The second is that echo-repetition questions will have different intonational characteristics from echo-surprise questions.

3.1 Participants

Data was collected in summer 2015 at the Phonetics Laboratory of the University of Deusto in Bilbao, Spain. A total of 22 native Spanish speakers from the Basque Country were recorded; their age range was 21–24. All were remunerated for their participation. Here we report on data from four female participants with clear Spanish dominance. Table 1 reports their age, city of residence, and their Bilingual Linguistic Profile scores in Spanish and Basque. These scores are calculated taking into consideration the history, use, proficiency and attitudes for each of these languages (BLP; Birdsong et al. 2012). In addition, the ‘dominance score’ reports the difference between the Spanish and Basque scores. Positive dominance scores indicate dominance in Spanish; negative scores indicate dominance in Basque. A score of or around 0 indicates balanced Spanish/Basque bilingualism.¹¹

Table 1. Participants

Participant	Spanish score	Basque score	Dominance score	Age	Town
AV	201	125	76	21	Leioa
NA	177	97	80	22	Bilbao
MM	199	31	168	21	Santurtxi
MA	161	76	85	22	Galdakao

3.2 Tasks and data collection

Participants were asked to complete a consent form and the BLP prior to being interviewed. The experimental materials comprised reading and elicitation tasks presented via PowerPoint in a laptop. Participants were provided with a series of different naturalistic contexts, with corresponding pictures and accompanying voiceovers by one of the authors. They listened to the audio via headphones and responded with either a question or statement, as appropriate. A short break was taken between tasks. The experiment was conducted in Spanish and lasted approximately one hour.

11. At the time that the experiment took place, there was no Basque BLP questionnaire available (there is one now). We created our own Basque questionnaire, modeled on the Spanish one, and written in Spanish as well.

Each task included four experimental sets; (1) fronted wh-questions (5 items), (2) information-seeking wh-in-situ questions (15 items), (3) echo-repetition questions (10 items) and (4) echo-surprise questions (10 items). The first set was constructed as a single question/answer context for comparison and will not be discussed further in this paper. In the other sets, target sentences were separated by statements or yes-no questions.

Because of the relative lack of frequency of wh-in-situ questions in Spanish (particularly with information-seeking readings; Reglero 2004), participants completed the reading task first, and all experimental blocks began with a few practice sentences. If participants did not produce wh-in-situ questions at the beginning of the elicitation task, the interviewer paused the recording and went through additional practice items with the participant before resuming the recording.

3.3 Data analysis

Data was recorded using a Tascam DR-05 portable digital recorder with two built-in omni-directional microphones, in mono at 44,000 Hz. Data analysis was done in *Praat* (Boersma and Weenink 2014) following Spanish ToBi labelling conventions (Aguilar et al. 2009; see also Hualde 2003; Face and Prieto 2007).

For this first phase of the project, we focus on the analysis of wh-in-situ questions from the elicitation task. We analyzed 10 information-seeking, echo-repetition, and echo-surprise questions for each speaker. For the first set, we chose items with wh-words comparable to those in echo-repetition and echo-surprise sets. The wh-questions elicited included *cómo* ‘how’, (*a*)*dónde* ‘where (to)’, *cuándo* ‘when’, *qué* ‘what’, and *quién* ‘who’. The total number of target sentences was 120 (10 sentences x 3 pragmatic readings x 4 participants). Five of these sentences were discarded because of errors or waveform distortions, leaving a total of 115 analyzed sentences.

As discussed in Section 2.3, impressionistic reports on the intonation of information-seeking vs. echo questions mention differences in (1) the final intonational contour of the sentence, and (2) the degree of stress on the wh-word. To test the former, we coded each elicited sentence for the location of the first prenuclear Low (L) tone, as well as the lowest Low and highest High (H) tone of the nuclear configuration. With these measurements, we calculated the *Global tonal range* and the *Local tonal range* (Willis 2004):

- (12) *Global tonal range*: The difference in Hz between the lowest Low tone of the first pre-nuclear accent and the highest High of the final boundary.
- (13) *Local tonal range*: The difference in Hz between the lowest Low tone in the nuclear configuration and the highest High of the final boundary.

We also examined the final boundary to ascertain whether the sentence had a rising (HH%) or falling (HL%) contour. Since both were attested in all conditions, we also calculated the *percentage of HH% final boundary contours* for each question type.

Figure 1a and 1b below exemplify how these measurements were taken. The first L coded in the tone tier is the first prenuclear Low; the second L is the nuclear Low, and H is the nuclear High. The final pitch value of each utterance (F) was also coded, providing an indication of whether the final contour was rising (as in 1a, where $H = F$), or falling (as in 1b, where $F < H$).

To test for stress differences on the wh-word, we measured its duration and the relative pitch of the High tone associated with it (pitch and duration are the main correlates of stress in Spanish; see for example Martínez-Celdrán and Fernández Planas 2007: 199–200 and references therein). We measured duration separately for each wh-word since they had different syllabic structures. In addition, we calculated the *duration ratio*, i.e., the duration of the wh-in-situ question relative to the duration of the sentence, to normalize durational values across wh-words, sentences, and speakers.

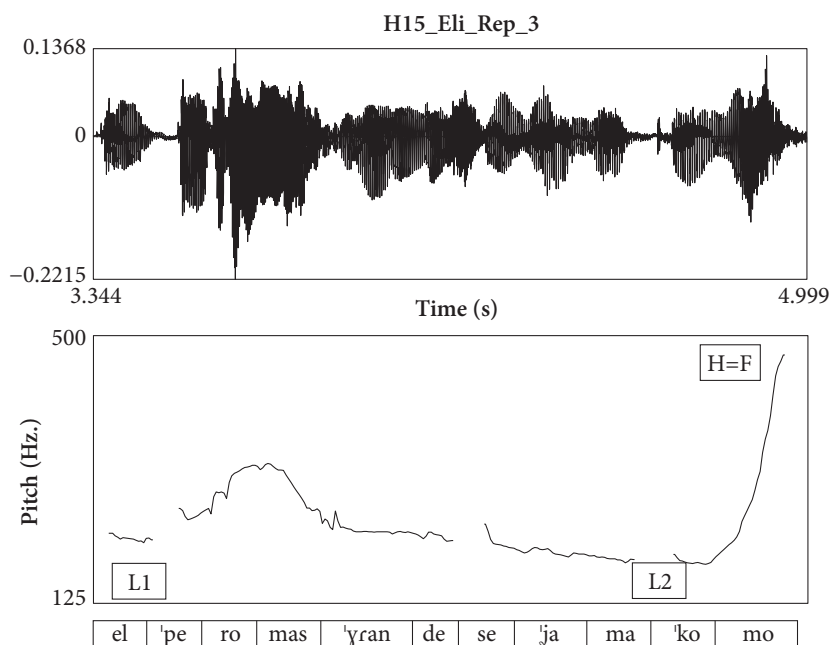


Figure 1a. Rising contour: ‘The biggest dog is called WHAT?’

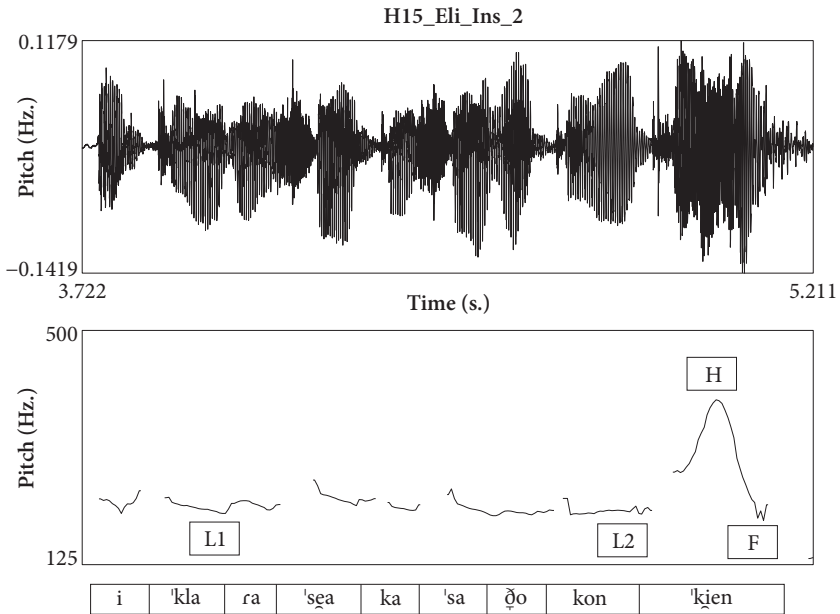


Figure 1b. Falling contour: ‘And Clara married who?’

4. Results

4.1 Final contours

An examination of intonational contours shows that rising (HH%) and falling (HL% or HM%) final contours are present in the three types of questions investigated, albeit with differing frequencies (Figure 2).¹² Rising contours are prevalent in information-seeking questions (in 33/39 of the tokens, or 85% of the time), while falling contours are most common in echo contexts (repetition: 23/39 or 60%; surprise: 21/38 or 55%). These findings contradict previous reports that information-seeking in-situ questions have falling intonation whereas echo-questions are predominantly rising (Uribe-Etxebarria 2002). On the other hand, these results offer strong support for our hypothesis that information-seeking and echo-questions are intonationally different, but do not provide evidence for intonational differences between echo-repetition and echo-surprise questions.

12. Since this is a preliminary analysis, we do not report the occurrence of finer-grained intonational contours such as circumflex (i.e., rising-falling) patterns, characterized here as falling. We also leave statistical analysis for a later stage of the project when more participants are analyzed. Note that Robles-Puente (2011) also found M% as a possible falling contour in Bilbao Spanish, which he attributes to truncation in sentences ending in an oxitone.

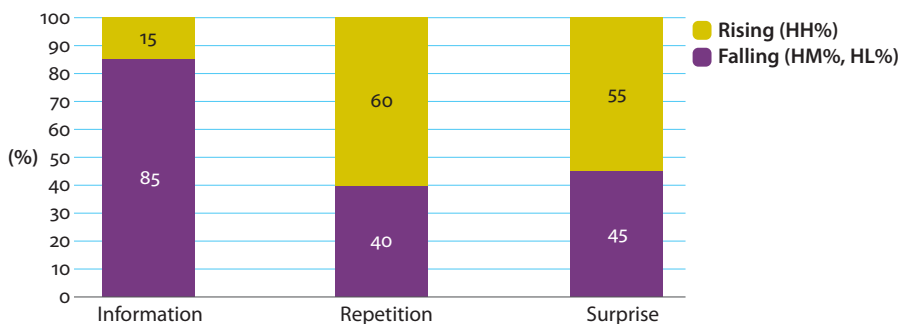


Figure 2. Final contours (speakers pooled; % rounded)

4.2 Tonal range

The comparison of Global and Local tonal ranges (GTR and LTR, respectively) shows that surprise questions have the steepest intonational slope, and information-seeking ones the flattest (Figure 3a, 4a). The average tonal ranges obtained and their equivalence in semitones are also provided in Figure 3b, 4b. Whereas the pitch values for the first prenuclear and the nuclear Low are relatively similar in all question types, the relative pitch of the nuclear High diverges more widely, being highest for surprise and lowest for information-seeking contexts.¹³ Thus, the nuclear High appears to drive the slope differences found among the three types of wh-questions. These findings are consistent with previous reports that echo-questions have ‘stronger’ or ‘sharper’ intonation than information-seeking questions. Since pitch is also a correlate of stress, these results are also compatible with stronger/ more marked stress on the wh-word in echo-questions.

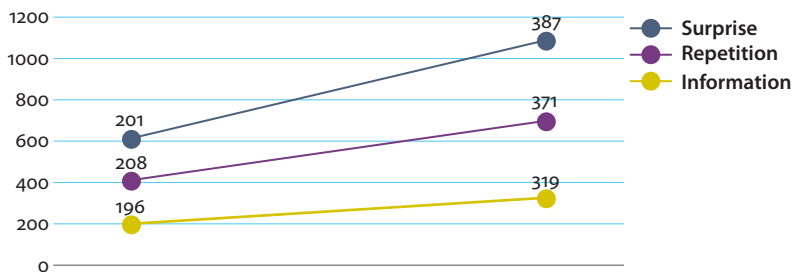


Figure 3a. Global Tonal Range

13. Global tonal ranges are also somewhat lower than Local tonal ranges; the reason is that the nuclear Low is typically lower in pitch than the first prenuclear Low.

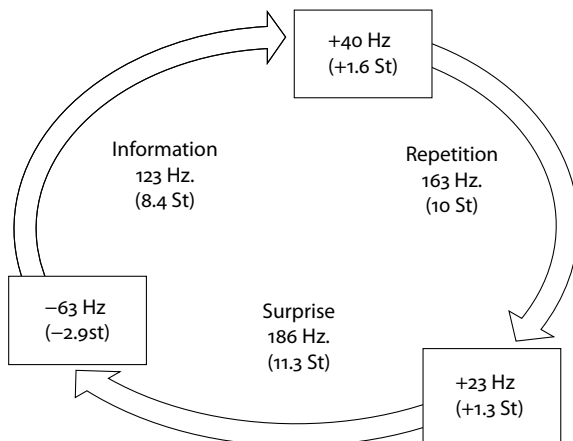


Figure 3b. Global Tonal Range differences

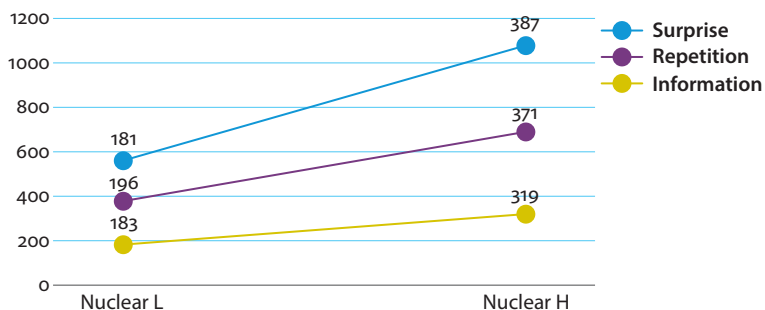


Figure 4a. Local Tonal Range

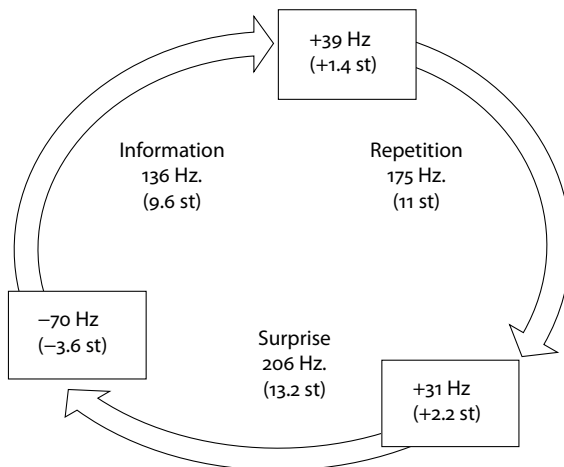


Figure 4b. Local Tonal Range differences

Figure 3b, 4b display schematically in Hz. and semitones the relative tonal differences among the three types of questions examined. The first generalization that emerges is that the GTR for echo questions is on average higher than that of information-seeking questions (cf. also Navarro Tomás 1944). The GTR of information-seeking questions is perceptually different from that of echo contexts (Figure 3b) – for both, it surpasses the perceptual threshold of 1.5 st. (cf. 't Hart 1981). This supports the hypothesis that information-seeking questions differ intonationally from echo-questions. Figure 4b shows that the LTR of surprise questions is perceptually distinct from that of repetition and information-seeking contexts, supporting the hypothesis that echo-repetition and echo-surprise questions are different intonationally.

4.3 Duration

Table 3 shows the average duration of wh-questions in the three pragmatic contexts investigated, as well as the corresponding duration ratio (i.e., the duration of the wh-in-situ question relative to the duration of the sentence). We will focus on the latter, since it normalizes durational measurements across wh-words, sentences and participants. Results show large variability; the durational ratio is highest in information-seeking contexts for *cuándo* 'when' and *quién* 'who'; in repetition contexts, for *adónde* 'where', and in surprise contexts, for *cómo* 'how'.¹⁴ For *qué*, duration ratio differences across contexts are negligible.

When the durational ratios are averaged across the wh-words examined, information-seeking contexts appear to have the highest durational ratio, and surprise contexts the lowest. Since duration is one of the correlates of stress, these findings

Table 3. Duration

	Information-seeking		Repetition		Surprise	
	wh-word	ratio	wh-word	ratio	wh-word	ratio
<i>adónde</i> 'where' (12 tokens)	366 ms.	32%	439 ms.	35%	376 ms	27%
<i>cómo</i> 'how' (12 tokens)	326 ms.	28%	339 ms.	22%	374 ms.	36%
<i>cuándo</i> 'when' (8 tokens)	393 ms.	30%	403 ms.	26%	–	–
<i>qué</i> 'what' (40 tokens)	231 ms.	21%	240 ms.	22%	265 ms.	20%
<i>quién</i> 'who' (40 tokens)	357 ms.	27%	315 ms.	21%	339 ms.	22%
<i>Average</i>		27.6%		25.2%		21%

14. Note, however, that *cuándo* did not appear in surprise contexts, so durational results for this question are inconclusive.

provide modest support for the (variable) occurrence of ‘marked’ or ‘extra’ stress in the *wh*-question of information-seeking contexts, but not in surprise contexts.

5. Discussion and conclusion

This study investigates the intonational differences between information-seeking and echo contexts in Spanish *wh*-in-situ questions, and between echo-surprise and echo-repetition ones. Preliminary results from four Spanish-dominant female participants from the Basque Country show that information-seeking questions have a higher proportion of final rising contours than echo-questions.¹⁵ They also have a lower Global tonal range. Surprise and repetition echo-questions show a similar incidence of rising contours, but differ in that the former has a much higher Local tonal range than the latter. Thus, our analysis provides support for both hypotheses investigated.

The findings reported in this study do not confirm previous impressionistic reports that falling intonation characterizes information-seeking questions while rising intonation occurs mostly in surprise contexts (Uribe-Etxebarria 2002, Pope 1976), or that echo-surprise and echo-repetition questions have different intonational contours (cf. Escandell 1999). Our results also suggest that it is tonal range rather than the final contour that might lead into the perception of ‘strong’ or ‘upward’ intonation for echo-surprise questions (Pope 1976, Contreras 1999, Sobin 2010, Chernova 2013).

This paper also investigated possible correlates of ‘marked’ or ‘extra’ stress in information-seeking and echo-surprise questions; specifically, durational ratios and the relative pitch of the final High tone. It was found that, on average, information-seeking questions have the highest durational ratio, while the final High tone is highest in surprise contexts. Since duration and pitch are the main correlates of stress in Spanish, these findings suggest that both types of questions might be realized with ‘marked’ or ‘extra’ stress, realized (variably) as longer duration in the former, and (consistently) as higher pitch in the latter.

The present study is only the first phase in the investigation of the intonational and syntactic characteristics of Spanish *wh*-in-situ questions. Future research will expand the intonational analysis of the four participants reported here to the remaining 18. Research already under way compares the intonational properties of information-seeking *wh*-in-situ questions and comparable yes/no questions in NPS. Specifically, we discuss intonational evidence for their licensing by a null

15. Information-seeking questions also are reported to have final rising contours in French (Cheng and Rooryck 2000).

intonational morpheme, as proposed by Cheng and Rooryck (2000) for French. We also find evidence that the degree of language dominance impacts the intonational characteristics of these questions, particularly of the yes-no type, in line with Robles-Puente (2012) and Elordieta and Hualde (2014), a.o. (see González and Reglero to appear for more details).

We expect to extend this investigation to echo contexts, and to examine additional intonational and prosodic characteristics, including breaks (pauses) – reported before information-seeking wh-in situ words (Uribe-Etxebarria 2002) –, and their possible connection to focus.

Acknowledgements

Our deepest thanks to our research assistants Eric Martínez for assistance in the creation of the materials, and Jessica Craft for collecting the data and preparing it for acoustic analysis. We are indebted to Alex Iribar and the Phonetics Lab at the University of Deusto for kindly allowing us to use their facilities to collect data, and to Jon Franco for his help recruiting participants. We also thank Mark Amengual for assistance with the BLP, and Emma Ticio, Daniel Solís, Estrella Rodríguez and Vivianne Asturizaga for testing out the materials and suggesting improvements. Thanks also to Anel Brandl, Amy Bustin, Antje Muntendam, John Ohala, and the audiences at ASA 2015, LSRL 2016 and the FSU Spanish Graduate Student Conference in 2016 for their feedback and encouragement. Comments provided by two anonymous reviewers helped improve this paper. All shortcomings are of course ours. This research was funded by a COFRS grant awarded by the FSU Council on Research and Creativity to the second author in 2014–2015.

References

- Aguilar, Lourdes, Carme de-la-Mota, and Pilar Prieto (coords). 2009. *Sp_ToBI Training Materials*. http://prosodia.upf.edu/sp_tobi/
- Aoun, Joseph, and Yen-hui A. Li. 1993. “Wh-Elements in Situ: Syntax or LF?” *Linguistic Inquiry* 24: 199–238.
- Birdsong, David, Libby Gertken, and Mark Amengual. 2012. *Bilingual Language Profile: An Easy-to-Use Instrument to Assess Bilingualism*. <https://sites.la.utexas.edu/bilingual/>
- Boersma, Paul, and David Weenink. 2014. *Praat: doing phonetics by computer*. Version 6.0.15. <http://www.praat.org/>
- Bošković, Željko. 2000. “Sometimes in Spec, CP, Sometimes in Situ.” In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 53–87. Cambridge, MA: MIT Press.
- Bošković, Željko. 2002. “On Multiple Wh-Fronting.” *Linguistic Inquiry* 33: 351–383. <https://doi.org/10.1162/002438902760168536>
- Cable, Seth. 2010. *The Grammar of Q: Q-Particles, Wh-Movement and Pied-Piping*. Oxford England: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195392265.001.0001>

- Chang, Lisa. 1997. "Wh-in-Situ Phenomena in French." M.A. thesis, University of British Columbia.
- Cheng, Lisa Lai-shen, and Johan Rooryck. 2000. "Licensing Wh-in-Situ." *Syntax* 3: 1–19. <https://doi.org/10.1111/1467-9612.00022>
- Chernova, Ekaterina. 2013. "El análisis sintáctico de las preguntas *qu-in situ* de eco en español." *Verba* 40: 65–92.
- Chomsky, Noam. 1957. *Syntactic Structures*. Berlin: Mouton de Gruyter.
- Chomsky, Noam. 2000. "Minimalist Inquiries: The Framework." In *Step by Step: Essays on Minimalist Syntax in Honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and Juan Uriagereka, 89–155. Cambridge, MA: MIT Press.
- Cole, Peter, and Gabriella Hermon. 1998. "The Typology of Wh-Movement: Wh-Questions in Malay." *Syntax* 1: 221–258. <https://doi.org/10.1111/1467-9612.00009>
- Contreras, Heles. 1999. "Relaciones entre las construcciones interrogativas, exclamativas y relativas." In *Gramática descriptiva de la lengua española*, ed. by Ignacio Bosque, and Violeta Demonte, 1931–1963. Madrid: RAE-Espasa Calpe.
- Déprez, Viviane, Kristen Syrett, and Shigeto Kawahara. 2013. "The Interaction of Syntax, Prosody, and Discourse in Licensing French Wh-in-Situ Questions." *Lingua* 124: 4–19. <https://doi.org/10.1016/j.lingua.2012.03.002>
- Elordieta, Gorra, and José Ignacio Hualde. 2014. "Intonation in Basque." In *Prosodic Typology II: The Phonology of Intonation and Phrasing*, ed. by Sun-Ah Jun, 405–463. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199567300.003.0014>
- Escandell, Victoria. 1999. "Los enunciados interrogativos. Aspectos semánticos y pragmáticos." In *Gramática descriptiva de la lengua española*, ed. by Ignacio Bosque, and Violeta Demonte, 3929–3991. Madrid: RAE-Espasa Calpe.
- Etxepare, Ricardo, and Myriam Uribe-Etxebarria. 2005. "In Situ Wh-Phrases in Spanish: Locality and Quantification." *Reserches Linguistiques de Vincennes* 33: 9–34. <https://doi.org/10.4000/rlv.1238>
- Face, Timothy, and Pilar Prieto. 2007. "Rising Accents in Castilian Spanish: A Revision of Sp_ToBI." *Journal of Portuguese Linguistics* 6: 117–146.
- Ginzburg, Jonathan, and Ivan A. Sag. 2001. *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives*. Stanford, CA: CSLI.
- González, Carolina, and Lara Reglero. to appear. "The Licensing of Wh-in-Situ Questions: Intonational Evidence from Spanish." In *Selected Proceedings of the Hispanic Linguistics Symposium 2016*, ed. by Alfonso Morales-Front.
- Hualde, José Ignacio. 2003. "El modelo métrico-autosegmental." In *Teorías de la entonación*, ed. by Pilar Prieto, 155–184. Barcelona: Ariel.
- Huang, C. -T. James. 1982. "Logical Relations in Chinese and the Theory of Grammar." Ph.D. dissertation, Massachusetts Institute of Technology.
- Jiménez, María Luisa. 1997. "Semantic and Pragmatic Conditions on Word Order in Spanish." Ph.D. dissertation, Georgetown University.
- Kahnemuyipour, Arsalan. 2009. *The Syntax of Sentential Stress*. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199219230.001.0001>
- Lasnik, Howard, and Mamoru Saito. 1992. *Move-Alpha: Conditions on its Application and Output*. Cambridge, MA: MIT Press.
- Martínez-Celdrán, Eugenio, and Ana M. Fernández Planas. 2007. *Manual de fonética española: Articulaciones y sonidos del español*. Barcelona: Ariel.

- Mathieu, Eric. 2004. "The Mapping of Form and Interpretation: The Case of Optional Wh-Movement in French." *Lingua* 114: 1090–1132. <https://doi.org/10.1016/j.lingua.2003.07.002>
- May, Robert. 1985. *Logical Form: Its Structure and Derivation*. Cambridge, MA: MIT Press.
- Munaro, Nicola, and Hans-Georg Obenauer. 1999. "On Underspecified Wh-elements in Pseudo-Interrogatives." *University of Venice Working Papers in Linguistics* 9 (1–2): 181–253.
- Munaro, Nicola, and Hans-Georg Obenauer. 2002. "On the Semantic Widening of Underspecified Wh-Elements." In *Current Issues in Generative Grammar*, ed. by Manuel Leonetti, Olga Fernández Soriano, and Victoria Escandell, 165–194. Madrid: Universidad de Alcalá, UNED, and Universidad Autónoma de Madrid.
- Munaro, Nicola, Cecilia Poletto, and Jean-Yves Pollock. 2001. "Eppur si muove!: On Comparing French and Bellunese Wh-Movement." *Linguistic Variation Yearbook* 1: 147–180. <https://doi.org/10.1075/livy.1.07mun>
- Navarro Tomás, Tomás. 1944. *Manual de entonación española*. New York: Hispanic Institute in the United States.
- Obenauer, Hans-Georg. 2004. "Nonstandard Wh-questions and Alternative Checkers in Pagotto." In *Syntax and Semantics of the Left Periphery, Interface Explorations* 9, ed. by Horst Lohntein and Susanne Trissler, 343–384. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110912111.343>
- Obenauer, Hans-Georg. 2006. "Special Interrogatives – Left Periphery, Wh-Doubling, and (Apparently) Optional Elements." In *Romance Languages and Linguistic Theory 2004: Selected Papers from 'Going Romance'*, ed. by Jenny Doetjes and Paz González, 247–273. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.278.12obe>
- Pesetsky, David. 1982. "Paths and Categories." Ph.D. dissertation, Massachusetts Institute of Technology.
- Pires, Acrisio, and Heather Lee Taylor. 2007. "The Syntax of Wh-in-Situ and Common Ground." Paper presented at the 37th Linguistic Symposium on Romance Languages. University of Pittsburgh, March 15–18.
- Pope, Emily. 1976. *Questions and Answers in English*. The Hague: Mouton.
- Reglero, Lara. 2004. "A'-Dependencies in Spanish and Basque." Ph.D. dissertation, University of Connecticut.
- Reglero, Lara. 2007. "Wh-in-situ Interrogatives in Spanish." *Probus* 19: 267–297. <https://doi.org/10.1515/PROBUS.2007.009>
- Reglero, Lara, and Emma Ticio. 2013. "A Unified Analysis of Wh-in-Situ in Spanish." *The Linguistic Review* 30: 1–47. <https://doi.org/10.1515/trl-2013-0018>
- Robles-Puente, Sergio. 2011. "Absolute Questions Do Not Always Have a Rising Pattern: Evidence from Bilbao Spanish." In *Selected Proceedings of the 5th Conference on Laboratory Approaches to Romance Phonology*, ed. by Scott M. Alvord, 98–107. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, #2638.
- Robles-Puente, Sergio. 2012. "Two Languages, Two Intonations? Statements and Yes/No Questions in Spanish and Basque." *International Journal of Basque Linguistics and Philology (ASJU)* 46: 252–262.
- Sobin, Nicholas. 2010. "Echo Questions in the Minimalist Program." *Linguistic Inquiry* 41: 131–148. <https://doi.org/10.1162/ling.2010.41.1.131>
- Stjepanović, Sandra. 1999. "What do Second Position Cliticization, Scrambling and Multiple Wh-Fronting Have in Common?" Ph.D. dissertation, University of Connecticut.
- Stjepanović, Sandra. 2003. "A Word Order Paradox Resolved by Copy Deletion at PHON." *Linguistic Variation Yearbook* 3: 139–177. <https://doi.org/10.1075/livy.3.07stj>

- 't Hart, Johan. 1981. "Differential Sensitivity to Pitch Distance, Particularly in Speech." *Journal of the Acoustical Society of America* 69: 811–821. <https://doi.org/10.1121/1.385592>
- Uribe-Etxebarria, Myriam. 2002. "In Situ Questions and Masked Movement." *Linguistic Variation Yearbook* 2: 259–303.
- Vergnaud, Jean Roger, and María Luisa Zubizarreta. 2002. "Intervention Effects in the French Wh-in-Situ Construction: Syntax or Interpretation?" Ms., University of Southern California.
- Willis, Erik. 2004. "Dominican Spanish Absolute Interrogatives in Broad Focus." In *Laboratory Approaches to Spanish Phonology*, ed. by Timothy Face, 61–91. Berlin: Mouton de Gruyter.
- Zubizarreta, María Luisa. 1998. *Prosody, Focus, and Word Order*. Cambridge, MA: MIT Press.

Parametric comparison and dialect variation

Insights from Southern Italy

Cristina Guardiano,¹ Dimitris Michelioudakis,²
Guido Cordoni,² Monica Alexandrina Irimia,¹ Nina Radkevich²
and Ioanna Sitaridou³

¹Università di Modena e Reggio Emilia / ²University of York / ³University of Cambridge, Queen's College

This paper applies the Parametric Comparison Method (PCM) to the description of syntactic variation in the nominal domain in a representative subset of Romance dialects of Southern Italy. We observe that, in order to perform successfully at the level of micro-comparison, the method must be supplemented by parameters specifically targeting this level of resolution. We sketch some such (micro-)parameters, investigate their interaction with the observed surface patterns, and show that their distribution broadly matches the received wisdom about the dialectal structure of Italo-Romance.

Keywords: syntactic, microvariation, dialect classification, parametric comparison method, Southern Italo-Romance, noun phrase, adjectives, genitives, possessives, kinship nouns

1. Introduction

In this article, building on Guardiano et al. (2016), we apply the Parametric Comparison Method (PCM) (Guardiano and Longobardi 2005, Longobardi and Guardiano 2009)¹ to the contrastive analysis² of the following sample of Romance

1. For a description of the notion of parameter relevant for the PCM, and for a detailed description of the theoretical basis, the methodological procedures and the empirical testing of the method, the reader is also referred to Guardiano and Longobardi (2017) and Longobardi and Guardiano (2017).
2. All the data discussed in this paper come from the authors' own investigation: most of them have been collected specifically for the purposes of the present paper, others have been taken from our

dialects of Southern Italy (SIDs):³ three Upper Southern dialects (UpSDs) (Pellegrini 1977),⁴ namely Teramano (Teramo – TE),⁵ Campano (S. Maria Capua Vetere – CE) and Northern Calabrese (Verbicaro – CS, which in turn belongs to a conservative geo-linguistic zone known as the Lausberg area) (Lausberg 1939, Rohlfs 1966, Martino 1991), and five dialects from the Extreme Southern area (ExtSDs), namely Salentino (Cellino San Marco – BR), Southern Calabrese (Reggio Calabria – RC) and three dialects of Sicily (Mussomeli – CL, Ragusa – RG, Aidone – EN, the latter of remote Gallo-Italic origin).⁶

The structure of the paper is as follows. In Section 2, we discuss the syntactic classification of the SIDs of our sample. In Section 3, we describe the dimensions of microvariation observed in the sample, and discuss some parameters specifically targeting this level of resolution. Section 3.1 presents aspects related to adjectival modification, and in particular to the relative ordering of adjectives and the noun, a subdomain that shows high variability within Romance (Bernstein 1991, Crisma 1991, 1996, a.o.). In this domain, the SIDs are different from Italian and internally uniform, with the partial exception of Southern Calabrese. In the subdomain of ‘genitives’ (Section 3.2), a similar exception to the otherwise uniform behavior of the SIDs considered here is displayed by Northern Calabrese (Silvestri 2013), that retains the relics of a structure that was productive in Late Latin (Crisma and Gianollo 2006) and is also attested in Old Romance (Delfitto and Paradisi 2009). Section 3.3 introduces some aspects of the syntax of pronominal possessives, a domain quite variable across Romance. We explore a small subset of properties that breaks the unity of the Extreme Southern group and sets a neat areal divide between a ‘macro-Sicilian’ area (including the three dialects of Sicily and Southern Calabrese, but crucially not Salentino) and the rest. Finally, a property that unifies the ExtSDs of our sample, and sets a sharp separation with UpSDs, is the availability of implicit possessives with kinship nouns, described in Section 3.4. Section 4 sums up the main conclusions.

previous works. The data from currently spoken languages have been elicited from native speakers. All the examples of DPs cited in the paper have been tested as occurring in argument position.

3. See Map 1 in the Appendix.

4. For recent overviews of the classification of the dialects of Italy, see Maiden and Parry (1997), Loporcaro (2009), Ledgeway (2016).

5. Abbreviations: TE = Teramo; CE = (province of) Caserta; CS = (province of) Cosenza; BR = (province of) Brindisi; RC = Reggio Calabria; CL = (province of) Caltanissetta; RG = Ragusa; EN = (province of) Enna.

6. An overview of the history and sociolinguistic structure of these varieties (and the relevant literature) is provided in Guardiano (2014), Guardiano and Stavrou (2014), Guardiano et al. (2016).

2. A syntactic classification of Southern Italo-Romance

The parameters employed as *comparanda* in the PCM are selected on the basis of Longobardi's (2003) "Modularized Global Parametrization" strategy, i.e. aiming toward an eventually exhaustive systematization of syntactic variation within one specific module of grammar. The module used for our purposes is that of nominal syntax, as it may display a limited amount of interaction with other domains of syntax and a lesser degree of susceptibility to information-structure pressures which are often supposed to be triggers of diachronic instability.

Guardiano et al. (2016) showed that, in order to capture aspects of dialectal variation observed in a selected sample of Southern Italo-Romance dialects, a specific set of newly formulated parameters was needed; once added to the original parametric dataset, such parameters produced more precise taxonomic results. Here, we further pursue this line of research, and proceed to an exhaustive parametrization of the variation exhibited by the data.

In order to derive the observed differences across the dialects of our sample (that largely overlaps with the one used in Guardiano et al. 2016) in the most constrained and comprehensive way, six specific parameters were required, some of which are presented in Section 3, along with a short survey of the relevant data.

The tree in Figure 1 was generated from parametric pairwise Jaccard⁷ distances between 30 Indo-European languages belonging to six subgroups,⁸ including the eight SIDs of our sample.⁹ Such distances have been generated from 91 binary nominal parameters crucially including the aforementioned six (this dataset partially differs from the one used in Guardiano et al. 2016). The tree successfully retrieves most of the received wisdom about the articulation of Indo-European (cf. Longobardi et al. 2013), identifies Italo-Romance, and represents its dialectal classification in accordance with the received wisdom. The Extreme Southern unity is identified correctly: in this cluster, the two most (geographically) 'external' dialects

7. Cf. Guardiano and Longobardi (2017), Longobardi and Guardiano (2017) for details. The trees were produced through the software UPGMA. We calculated a consensus tree out of 100 datasets obtained through a bootstrapping procedure.

8. Languages and abbreviations: Far = Farsi; Ma = Marathi; Hi = Hindi; Pas = Pashto; E = English; D = German; Da = Danish; Nor = Norwegian; Rm = Romanian; Fr = French; Sp = Spanish; Ptg = Portuguese; It = Italian; Grk = (Standard Modern) Greek; CyG = Cypriot Greek; Blg = Bulgarian; SC = Serbo-Croat; Slo = Slovenian; Po = Polish; Rus = Russian; Ir = Irish; Wel = Welsh.

9. Abbreviations: Cam = Campano; Ter = Teramano; NCa = Northern Calabrese; Sal = Salentino; MsS = Mussomeli Sicilian; AdS = Aidone Sicilian; RGS = Ragusa Sicilian, SCa = Southern Calabrese.

of the sample (Mussomeli and Salentino) go together, as opposed to a more ‘central’ core (Ragusa, Aidone and Reggio Calabria). Northern Calabrese is the outlier of the Extreme cluster, a position that reflects its expected isolation from mainstream UpSDs. The two ‘central’ UpSDs of the sample are clustered together.

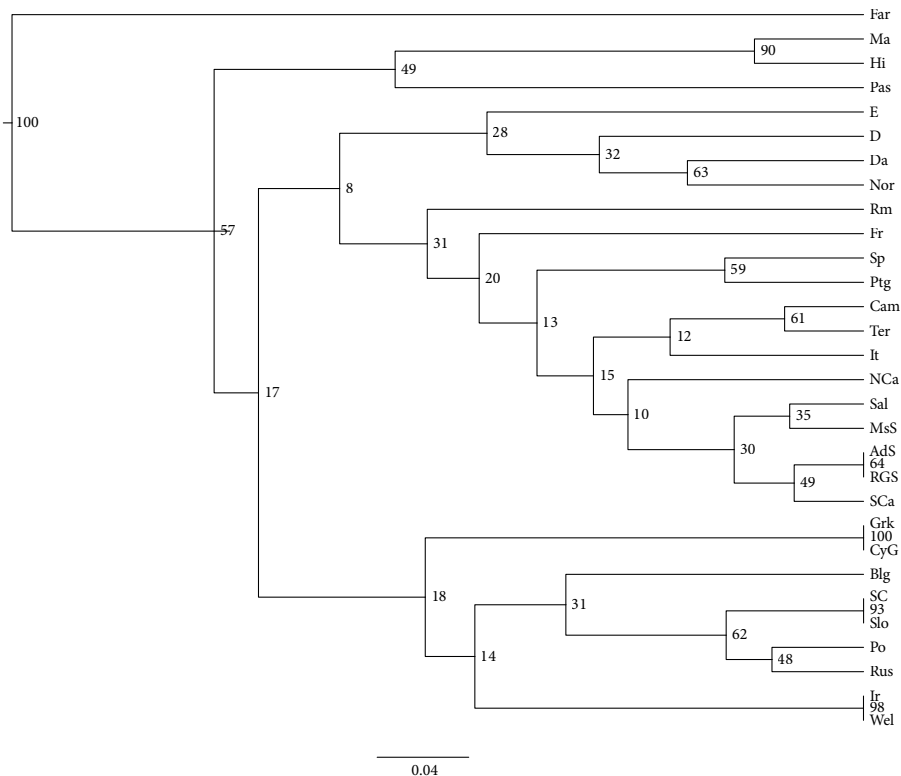


Figure 1. UPGMA tree from parametric Jaccard distances (91 parameters, IE languages and SIDs)

Expectedly, if we exclude the six aforementioned parameters, the taxonomic results remain generally correct, but are much less precise. The tree in Figure 2, generated from parametric pairwise Jaccard distances obtained from 85 parameters, correctly identifies the Italo-Romance group, but does not differentiate most dialects which are well distinguishable geographically, historically and phonologically/lexically (e.g. Pellegrini 1977).

These results are in line with the findings of Guardiano et al. (2016): the standard PCM practice of using taxonomic experiments to validate hypotheses of syntactic variability confirms that an accurate and ideally exhaustive parametrization of dialectal variation produces a faithful representation of what is independently known about the distribution and classification of the observed dialects.

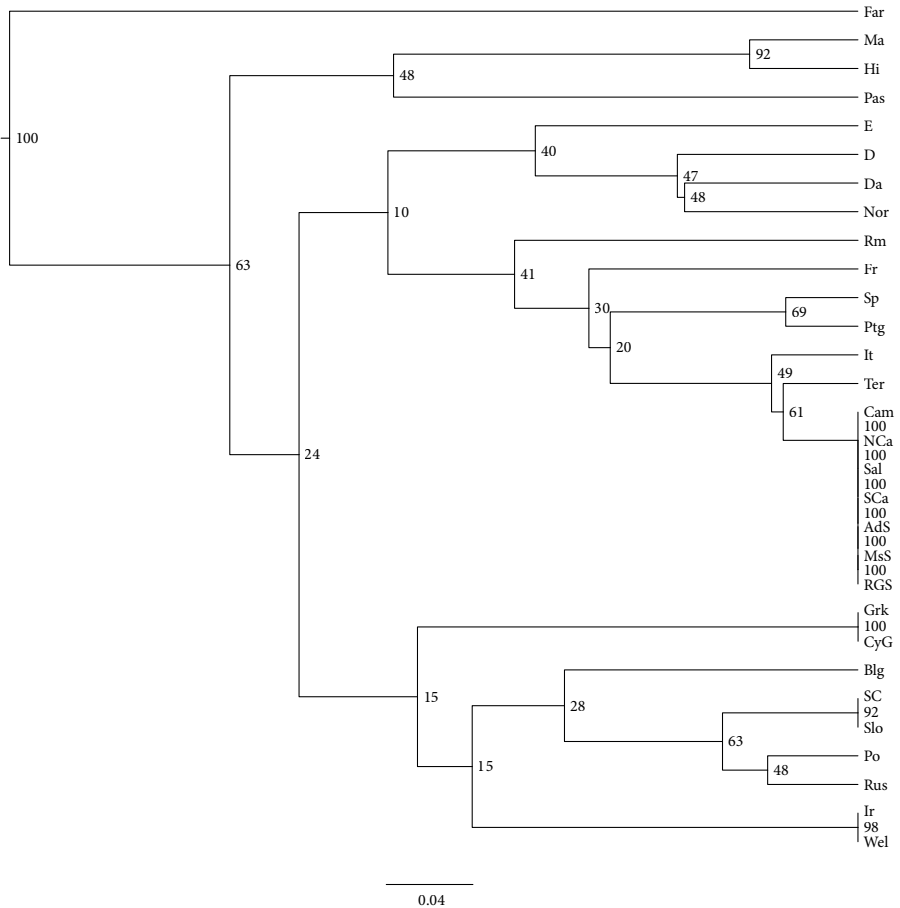


Figure 2. UPGMA tree from parametric Jaccard distances (85 parameters, Indo-European and the SIDs)

3. Towards higher resolution

The six parameters required in order to account for the specific aspects of variation observed across the Southern Italo-Romance dialects of our sample belong to the class of micro-parameters, i.e. parameters associated with intensionally definable classes of lexical items (Roberts 2012).

Two of them are taken from Guardiano et al. (2016), namely *Number on N* (FNN) and *Variable person in D* (FVP).¹⁰ The first (FNN) is connected to the

10. For a description of these parameters and their empirical correlates in the Romance dialects of Southern Italy, along with the relevant literature, see Guardiano et al. (2016). A summary of

representation of the feature Number on nouns, and its most visible syntactic effect is the possibility for a language to allow determinerless nouns in argument position. This parameter sets a neat divide between two broad areas in Central/Southern Italy. The first is represented by Campano and Teramano, which seem to behave like some other Central and Northern dialects of Italy, and notably in a way similar to French. The other includes the ExtSDs and Northern Calabrese, which instead behave like (Standard) Italian. The second parameter (FVP) presents a phenomenon known as “unagreement”, broadly connected to the representation of the feature ‘Person’ in D. This phenomenon, unlike most of the other syntactic properties related to the representation of Person in D, does not seem to provide any significant genealogical information, not even at the level of dialectal diversification: the distribution of the relevant parameter values is scattered, and non-reducible to any historical, geographical or sociolinguistic pattern.¹¹

The other four parameters result from a revision, expansion, and a novel systematization of the observed space of variation, and are briefly presented in the following sections.

3.1 Adjectives and noun movement

According to the literature (Alexiadou et al. 2007, for original references), some types of adjectives are first-merged prenominal, below determiners and numerals, and ordered according to a semantic hierarchy (the first one proposed by Sproat and Shih 1991; see also Scott 2002, Alexiadou et al. 2007, a.o.). The instantiation of such a hierarchy that we adopt here is shown in (1):

- (1) High [Subject-oriented] > Manner 1 [size, value, ...] > Manner 2 [color, shape, ...] > Argument [provenance, ...]

We label this type of adjectives ‘structured’ adjectives. In our system, the availability of structured adjectives in a language is encoded by the parameter $\pm structured$

the distribution of their values in the languages of our sample and of their empirical manifestations can be found in the Appendix of the present paper (Tables A–D).

11. The most ‘external’ (in terms of geographical position) dialects of the area considered here (i.e. Teramano, Campano, Mussomeli, Salentino) share the value (–), as opposed to a geographically more compact core that includes Eastern Sicily and Calabria. Notice also that this parameter signals the only difference observed in the nominal domain between the varieties of Sicily, that for the rest are parametrically uniform. Moreover, the inspection of its values outside Italo-Romance reveals a similar geography-driven distribution: (+) in Spanish and Catalan, though not in Portuguese or French, (+) in Greek, though not in all dialects, (+) in standard Arabic though not in Hebrew, etc. This situation confirms the impression that such values can at best provide areal rather than genealogical information.

A(adjective)s (AST), that has the value (+) in Germanic and Romance, (–) in many African languages (e.g., Wolof). The pre-/postnominal linearization of structured adjectives depends on the absence/presence/scope of N(oun)/NP-movement (Bernstein 1991, Crisma 1991, 1996, Laenzlinger 2005, Cinque 2010, a.o.). The different options of N-movement across structured adjectives are defined, in our system, by parameters $\pm N$ over *As* (NOA), $\pm N$ over *M2 As* (NM2), and $\pm N$ over *M1 As* (NM1) (listed in Table 1), which are only relevant in languages with structured adjectives (+AST).

Table 1. Structured adjectives and N-movement: all theoretically possible crosslinguistic combinations

Label	Parameter	Implicational conditions ¹²	I	II	III	IV
NOA	$\pm N$ over <i>As</i>	+AST	–	+	+	+
NM2	$\pm N$ over <i>M2 As</i>	+NOA	0	–	+	+
NM1	$\pm N$ over <i>M1 As</i>	+NM2	0	0	–	+
			English/Latin	Walloon ¹³	Italian ¹⁴	Celtic ¹⁵

Table 1 illustrates all the possibilities of crosslinguistic variation (4 types: I–IV) predicted by the three binary parameters. As shown below, SIDs instantiate type IV, namely a language where N/NP crosses over even the highest class(es) of adjectives.

The empirical manifestations which set the value (+) for the three parameters of Table 1 are listed in Table 2. Notice that setting the value (+), as opposed to (–),

12. In Longobardi and Guardiano's (2009) formalism, implicational conditions between parameters define cases in which the content of a parameter is entirely predictable on the basis of the setting of another parameter (i.e. the parameter is unsettingtable). In this paper, we present a simplified version of such cross-parametric implications.

13. (i) *one bèle bleuve cote alemande* (Bernstein 1991)
 a beautiful blue dress German
 'a beautiful blue German dress'

14. Italian (i), French, Spanish, Portuguese and Romanian all instantiate type III, with minor lexical differences:

(i) a. *un famoso vestito rosa francese* (, *indossato da Jacqueline Kennedy il 22 Novembre*)
 a famous dress pink French
 'a famous pink French dress (, worn by Jacqueline Kennedy on November 22)'
 b. **un famoso rosa vestito francese*
 c. **un famoso rosa francese vestito*

15. Cf. Rouveret (1994), a.o.

means adding a movement rule to the grammar of a given language. In our system, the value (+) of any parameter can be set if positive evidence is available, while the value (–) can be assigned by default.¹⁶ Notice also that, in the particular case of the three parameters of Tables 1 and 2, any of the conditions listed for each parameter is sufficient for the value (+) to be set. Finally, further language-specific features may occasionally interact with the empirical evidence relevant to set a given parameter, producing misleading surface structures which would set the wrong value. For instance, as indicated in Table 2, in many languages an N-movement parameter can be set thanks to the existence of strings containing postnominal adjectives in positions excluding other types of adjectives (i.e. non-structured/reduced relatives); such positions are between the noun and an adnominal argument (genitive/PP) or between the noun and a postnominal restrictive relative. Nevertheless, in other types of languages, such as for instance Greek, such orders are possible thanks to independent properties (e.g. predicational/appositive structures, reduced relative clauses, etc.) (Alexiadou et al. 2007, Alexiadou 2014, and references therein), while the grammar also allows prenominal placement of the respective adjectives. Thus, the kind of manifestation just described should be relativized accordingly, by specifying that it does not apply if [Adj N] orders are available. For cases like this, when required, we add to each manifestation a list of empirical conditions under which that manifestation does not apply (*‘elsewhere’ conditions*). In Table 2, we exemplify some such conditions.

In terms of our system of parameters and setting rules, Table 2 is representative of the typical format of manifestations. In principle, each manifestation consists of: (1) a (possibly null) set of conditions that must be met for the manifestation to apply (*initial conditions*); such conditions depend on crossparametric implications; (2) the actual manifestation(s), i.e. the list of empirical conditions that must be met for the value (+) to be set; (3) a (possibly null) set of conditions that, if met, don’t allow the manifestation(s) to apply (*‘elsewhere’ conditions*).

16. The symbol ‘0’ indicates that the parameter is not set as a consequence of crossparametric implications.

Table 2. Empirical manifestations of the parameters of Table 1

Parameter	Manifestation(s)	(+)
$\pm N$ over As	<i>Initial condition</i>	
	The language has structured adjectives. ¹⁷ [...]	
	<i>Manifestations</i>	
	In argument position:	
	a. Argument adjectives (e.g. those denoting <i>origin</i> , <i>nationality</i> , etc.) follow the noun in discourse-neutral contexts.	YES
	b. If a language has a non-adpositional structured genitive occurring after structured adjectives (GenO), postnominal Argument adjectives occur before GenO.	YES
	c. Postnominal Argument adjectives occur before prepositional phrases.	YES
	d. Postnominal Argument adjectives occur before relative clauses.	YES
	<i>'Elsewhere' conditions</i>	
	The manifestations do not apply if the order [ArgAdj N] is available (in unmarked contexts). In languages with postnominal adjectives and with 'polydefinite' structures, postnominal APs in indefinite nominals do not count to set the value (+). [...]	
$\pm N$ over M2 As	<i>Initial condition</i>	
	The noun raises over Argument adjectives.	
	<i>Manifestations</i>	
	In argument position:	
	a. M2 adjectives (e.g. <i>color</i>) [and Argument adjectives (e.g. those denoting <i>origin</i> , <i>nationality</i> , etc.)] follow the noun in discourse-neutral contexts.	YES
	b. If a language has a non-adpositional structured genitive occurring after structured adjectives (GenO), postnominal M2 [and Argument] adjectives occur before GenO.	YES
	c. Postnominal M2 [and Argument] adjectives occur before prepositional phrases.	YES
	d. Postnominal M2 [and Argument] adjectives occur before relative clauses.	YES

(continued)

17. Namely, the parameter AST (cf. above) must have the value (+) in order for this (and the following two) parameter to be actually set. Otherwise they have a 0. Notice that further initial/elsewhere conditions apply to the manifestations of this particular parameter, which are not listed here because irrelevant for the languages under analysis.

Table 2. (continued)

Parameter	Manifestation(s)	(+)
	<i>'Elsewhere' conditions</i> The manifestations do not apply if the order [M2Adj N] is available (in unmarked contexts). In languages with postnominal adjectives and with 'polydefinite' structures, postnominal APs in indefinite nominals do not count to set the value (+). Manifestation (a): if a language allows multiple postnominal adjectives, belonging to different classes (e.g. Speaker-oriented and M2) which can be ordered freely in postnominal position, this evidence does not set the value (+). [...]	
$\pm N$ over M1 As	<i>Initial condition</i> The noun raises over (Argument and) M2 adjectives. <i>Manifestations</i> In argument position:	
	a. M1 adjectives (e.g. <i>size</i>) [M2 adjectives (e.g. <i>color</i>) and Argument adjectives (e.g. those denoting <i>origin</i> , <i>nationality</i> , etc.)] follow the noun in discourse-neutral contexts.	YES
	b. If a language has a non-adpositional structured genitive occurring after structured adjectives (GenO), postnominal M1 [M2 and Argument] adjectives occur before GenO.	YES
	c. Postnominal M1 [M2 and Argument] adjectives occur before prepositional phrases.	YES
	d. Postnominal M1 [M2 and Argument] adjectives occur before relative clauses.	YES
	<i>'Elsewhere' conditions</i> The manifestations do not apply if the order [M1Adj N] is available (in unmarked contexts). In languages with postnominal adjectives and with 'polydefinite' structures, postnominal APs in indefinite nominals do not count to set the value (+). Manifestation (a): if a language allows multiple postnominal adjectives, belonging to different classes (e.g. Speaker-oriented and Manner 1), which can be ordered freely in postnominal position, this evidence does not set the value (+). [...]	

The SIDs explored here exhibit positive evidence to assign a value (+) to each of the three parameters. Indeed, Argument, M2 and M1 adjectives follow the noun, as shown by the Ragusa examples in (2), (3) and (4), respectively.

- (2) a. *na màkina tedèska ka mi pjàci /kku ccinku sportèlla*
a car German that to.me like.3s/with five doors
'a German car that I like / with five doors'

- b. **na tedèska màkina*
a German car
- (3) a. *na màkina bblù ka mi pjàci /kku ccìнку ṣportèlla*
a car blue that to.me like.3s/ with five doors
'a blue car that I like / with five doors'
- b. **na bblù màkina*
a blue car
- (4) a. *na màkina nòva/ròssa ka mi pjàci /kku ccìнку ṣportèlla*
a car new/big that to.me like.3s/ with five doors
'a new car that I like / with five doors'
- b. **na nòva/ròssa màkina*
a new/big car

Exactly as in Celtic (an example of Type IV in Table 1) (Rouveret 1994), there are only few lexically specified exceptions: only the *speaker-oriented* counterparts of *bello* 'beautiful, nice' (5)¹⁸ are allowed in prenominal position.

- (5) a. *nu bbèllu vagnòne* [Salentino]
a nice boy
'a handsome boy'
- b. *nu bbèllu swònnu* [Ragusa]
a nice sleep
'a good sleep'

Thus, in this particular domain, the SIDs are uniform, and represent for Romance the empirical instantiation of a type of language that follows from the settings of an already existing and independently motivated group of parameters (+NOA, +NM2, +NM1): Type IV in Table 1.

Southern Calabrese displays a peculiarity not shared by any other dialect of our sample: it accepts (some) M1 adjectives in prenominal position when they receive emphatic interpretation (6b), a possibility that is not predicted by any combination of parameter values in Table 1.

- (6) *Paskàli àvi na GRÀNDI/*gràndi kàsa (gràndi)*
Pasquale has a big house (big)
'Pasquale owns a really big house'
(the meaning *a house that is big* is only available if *gràndi* is postnominal)

18. And/or of *buono* 'good', more rarely of *brutto* 'ugly' and *cattivo* 'bad', and a few others, varying across the dialects (Guardiano and Stavrou 2014, in prep.; Guardiano et al. 2016).

The availability of an information-structure related position that (optionally) attracts high structured adjectives above their canonical position is encoded by the parameter \pm *Emphatic A fronting* (EAF), that can only be set in languages where N crosses high structured adjectives (hence the implicational condition +NM1), and has the value (+) in Southern Calabrese, (-) in the rest of SIDs, as shown in Table 3.

Table 3. Emphatic A(djective) fronting

Label	Parameter	Impl. cond.	It, Wal, Lat	Sicily, NCa, Cam, Ter	SCa
EAF	\pm <i>Emphatic A fronting</i>	+NM1	0	-	+

The empirical manifestation(s) to set the value (+) of the parameter EAF are listed in Table 4.

Table 4. Manifestation for the parameter Emphatic A(djective) fronting

\pm <i>Emphatic A fronting</i>	(+)
<i>Initial condition</i>	
M1 adjectives are postnominal (i.e. crossed over by N) in unmarked configurations.	
<i>Manifestation(s)</i>	
Some M1 adjectives can be prenominal when emphatic.	YES

3.2 Genitives

By the label ‘genitive’ we refer to direct adnominal arguments. Genitives can be divided into two classes: (i) free and (ii) functional (Longobardi and Silvestri 2013). Free genitives are freely iterable and can be formally realized with adpositions or inflectional marking. If the genitive is not iterable, we are dealing with ‘functional’ genitives, i.e. genitives occupying fixed functional positions. We identify two positions for functional genitives: GenS and GenO.¹⁹ The former precedes all structured adjectives, while the latter follows them, as in (7), abstracting away from the surface position of N.

(7) D > GenS > Manner 1 As > Manner 2 As > Argument As > GenO > N

In Latin, all the positions crosslinguistically available for the nominal (i.e. non-clausal and non-adpositional) arguments of a noun (i.e. GenS, GenO, FreeGen) were active. Contemporary standard Romance languages, instead, with the

19. The two functional genitive positions are accessible only to the first argument of a strict thematic hierarchy (Possessor>Agent>Theme). Functional genitives are never adpositional, and must meet particular licensing conditions: either phi-agreement between a head noun and a genitive or overt/covert N-movement over the genitives. See Longobardi et al. (2016) for details.

exception of Romanian,²⁰ do not display any trace of non-adpositional functional genitives.²¹

However, various works have pointed out instances of a genitive with the characteristics of GenO (i.e., postadjectival, non-iterable and non-prepositional) in Old Romance, e.g. in Old French (*li filz Marie*, ‘Mary’s son’, *li chevaus le roi*, ‘the king’s horse’, *la meson son pere*, ‘his father’s house’, ...) (Foulet 1919, Arteaga 1995, Delfitto and Paradisi 2009, Gianollo 2009, Sitaridou 2017, a.o.). Yet, unlike in languages where GenO is productive, in Old Romance this non-prepositional genitive was subject to restrictions concerning the type of head noun and the semantic relationship between the head and the genitival DP (only nouns denoting body parts, toponyms, kinship terms, inalienably possessed or so perceived items such as *home*, *land*, etc. seem to accept it). The parameter \pm *partial GenO* (PGO, Table 5) distinguishes this type of lexically conditioned genitive from the fully productive GenO.

Table 5. Partial GenO

		Sp, Fr, Ptg, It, Sicily, SCa, Sal, Cam, Ter	NCa, Old Romance
PGO	\pm <i>partial GenO</i>	–GFO ²²	–
			+

The empirical manifestations which set the value (+) of parameter PGO are listed in Table 6. Notice that the value (+) can only be set if *all* the conditions listed in Table 6 are met. The value (–), instead, can be assigned by *default*.

Table 6. Manifestations of Partial GenO

\pm <i>partial GenO</i>	(+)
<i>Initial condition</i>	
No systematic instances of post-adjectival non-prepositional genitives.	
<i>Manifestation(s)</i>	
A non-adpositional genitive occurs after (all types of) structured adjectives.	
A non-adpositional genitive occurs with a head noun that belongs to any of the following classes: body parts, toponyms, kinship terms, inalienably possessed items, <i>home</i> , <i>land</i> .	
	YES

20. As shown by Irimia et al. (in prep), Romanian is likely to activate the GenS position.

21. Apart from possessive pronouns, cf. Section 3.3.

22. Parameter GFO (\pm GenO) asks if a language checks a non-adpositional genitive in a functional postadjectival position (+GFO) or not (–GFO). In + GFO languages, all types of argument nouns can go in GenO and all types of head nouns can license GenO. Therefore, in this type of languages, PGO is not set. Since Romance languages are all –GFO, parameter PGO is relevant (i.e. it can be set).

In Southern Italy, positive evidence for the value (+) of this parameter is visible nowadays in the Lausberg area (8) (Silvestri 2013, and subsequent work).

- (8) a *buttighja gròssa u vinu* [Northern Calabrese]
 the bottle big the vine
 ‘the big wine bottle’

The other SIDs feature lexicalized residuals of this construction, e.g. in some fixed expressions, along with a few toponyms, still in use in Ragusa (9) (Guardiano 2014).

- (9) a. *pp' àmùr(i) u Signùri*²³
 for the.love the Lord
 Lit. ‘for God’s love’
 b. *Gjànn(i) u ggwàrdia*
 Gianni the cop
 ‘John, the cop’s son’ (lit. ‘John of the cop’)

Cases like those listed in (9), also found in other dialects considered here, suggest that a GenO was reasonably productive in older stages of these dialects, presumably until recent times, but is not productive anymore in the current languages. Therefore, in all SIDs but Northern Calabrese PGO is assigned the value (–).

3.3 Possessives

By the label ‘possessive’ we refer to pronominal forms fulfilling a ‘genitive’ function.²⁴ As a first approximation, three types of possessives are attested across Romance: (i) ‘definiteness-checking’ (attested in Spanish or French), that are incompatible with e.g. definite/indefinite articles, demonstratives, quantifiers etc., and carry a definite meaning (parameter PDC in Table 7); (ii) ‘definiteness-matching’, that can only be licensed in adjacency to a definite affix on the head noun, and available in Romanian (parameter DMP in Table 7; DMP is neutralized in the other Romance languages for lack of the triggering context) (Irimia et al. in prep.); (iii) ‘adjectival’, that in argument position must co-occur with articles, demonstratives or quantifiers, and do not carry definite meaning by themselves or require to be licensed by a relation with a definite element, as in Portuguese and Old Spanish (parameter APO in Table 7).

23. This alternates with the prepositional variant *pp' àmùri r'o Signùri* (lit. ‘for the love of the Lord’).

24. As such, they relate to all theta-roles realized by genitives, namely P(ossessors), S(ubjects), O(bjects).

Table 7. Possessives

	Parameter	Implications	Latin	Spanish	French	Romanian	Italian
PDC	\pm <i>def checking possessives</i>	+ <i>article</i>	0	+	+	-	-
DMP	\pm <i>def matching possessives</i>	+ <i>enclitic article</i> ²⁵	0	0	0	+	0
APO	\pm <i>adjectival possessives</i>		+	+	- ²⁶	-	+

All the SIDs considered here are uniformly -PDC, 0DMP and +APO, like Italian. They all display at least one class of adjectival possessives occurring postnominally and (unlike Italian) not necessarily informationally marked in that position, as shown in (10) and (11).²⁷

- (10) a. *a màkina mìà* [Ragusa]
the-F.S car-F.S my-F.S
‘my car’
b. *i màkini mjèi*
the-P car-F.P my-F.P
‘my cars’
c. i. **a mìà màkina*
ii. **i mjèi màkini*
- (11) a. *lu libbru tòà* [Salentino]
the book your-s
‘your book’
b. *li libbri tòi*
the-P book-P your.P
‘your books’

25. Outside Romance, this parameter is set to the value (+) in Bulgarian, Norwegian and Icelandic, to the value (-) in Danish.

26. Cf. fn. 38 below.

27. Postnominal possessives are overtly inflected in Sicilian and Southern Calabrese (where they agree in gender and number with the head noun), in Campano and Salentino (where they agree in number with the head noun), uninflected in Northern Calabrese and Teramano. We leave this morphological variation outside the parametric picture because so far it seems to be syntactically inconsequential. In Teramano, postnominal possessives display further peculiarities (for instance, they require, under specific conditions, an epenthetic [-a] that attaches to the noun and is also required with some types of postnominal adjectives) which are not encoded in our parametrization either: cf. Mantenido (2015) and references therein.

- c. i. **lu tòa libbru*
 ii. **li tòi libbri*

In Campano, Teramano, Northern Calabrese and Salentino this is the only type of possessives available generally.²⁸ The dialects of Sicily, instead, have a further form of ‘adjectival’ possessives besides postnominal ones: a class of prenominal items, with reduced morphology (i.e., no gender and number agreement with the head noun), and obligatorily co-occurring with a D-item (definite or indefinite), as shown in (12).²⁹

- (12) a. i. *a mo³⁰ màkina* [Ragusa]
 the-F.S my-F.S car-F.S
 ‘my car’
 ii. *na mo màkina*
 a-F.S my-F.S car-F.S
 ‘a car of mine’
 b. **a / na màkina mo*
 c. **mo màkina*

Southern Calabrese is similar in the relevant respects, with some minor details to be addressed.³¹

Giorgi and Longobardi (1991, Chapter 3, Appendix) noted a further distinction between weak and strong surface positions for possessives (as well as some

28. In these dialects, a further type of possessives, uninflected and enclitizing onto the head noun, is actually available (e.g. Verbicaro *patrà-ma* ‘father-my’, Silvestri 2013: 116–119; cf. also Guardiano 2014: 91–93 and, for further data, recent discussion and the literature, Manzini and Savoia 2007, D’Alessandro and Migliori 2013). Yet, their distribution is constrained. Indeed, such possessives are licensed only under the following conditions: (i) they only attach to certain types of head nouns (kinship nouns and a few others, a class widely attested crosslinguistically: cf. WALS, van Rijin 2016, etc.) in the singular; (ii) they never co-occur with an article; (iii) no modifier precedes the noun (Silvestri 2013: 132). This is a typical diagnostic for N raising to D in Romance. Such constructions seem to be a counterpart of the standard Italian construction in which, in a maximality interpretation, a singular kinship noun with a prenominal possessive must raise to D along with the possessive itself (Longobardi 1996), thus dispensing with the definite article.

29. A similar distribution is also visible in some Central and Northern dialects (e.g., Toscano and Padovano, Giorgi and Longobardi 1991; cf. also Manzini and Savoia 2007: 553–557, Cardinaletti 1998, and literature therein). See also Old Italian (Salvi and Renzi 2010, Chapter 9 and 40).

30. Some varieties have the form *ma*, equally invariable for gender and number.

31. In Southern Calabrese, prenominal possessives (i) are uninflected: like in Sicily, these possessives cannot be contrastive, require an article (definite or indefinite) (ia–b), precede numerals and are never found in postnominal or non-adnominal position (ic–d).

other adjectives and determiners): the weak forms are in proclisis to the head noun or prenominal adjectives; the strong ones fail this condition, and typically occur postnominally or with a null noun.³² Specific adjectival possessives can be strong only (e.g. Sp. *mío*), weak only (Paduan *me*), or ambiguous (Italian *mio*). Given the observations above, SIDs and (Standard) Italian seem to instantiate three types of languages:

-
- (i) a. *lu/nu mè libbru*
 the-M.S/a-M.S my book-M.S
 'my book / a book of mine'
 b. **mè libbru*
 c. **lu/nu libbru mè*
 d. *(*ki*)*stu libbru èste (lu) mè*
 this-M.S book-M.S is the-M.S my
 Intended: 'this book is mine'

Fully agreeing possessives are normally used in non-adnominal and non-argument function (ii).

- (ii) a. (*ki*)*stu libbru èste (lu) mèu*
 this-M.S book-M.S is the-M.S my-M.S
 'the book is mine'
 b. *lu mèu è kistu*
 the-M.S my-M.S is this-M.S
 'mine is this one'

They are marginally accepted in prenominal position, with strongly emphatic reading (iii). The possibility of fronting adjectival possessives under informationally marked readings is expected from the parameter +*fronted prenominal As*.

- (iii) [?]*lu MÈU libbru*
 a-M.S my-M.S book-M.S
 'my book' (marginally accepted, only if *mèu* is emphatic)

Finally, fully inflected possessives are accepted (though less frequent than the 'reduced' prenominal ones) in postnominal position (iv). For further data, cf. for instance Manzini and Savoia (2007: 557, 563).

- (iv) *lu libbru mèu*
 the-M.S book-M.S my-M.S
 'my book'

32. If there is a formal distinction between the two types of items, it may be the case that a strong form has more morphological features than a weak one (Spanish *mío(-s)/mía(-s)* vs. *mi(-s)*), never the opposite.

- (13) a. Languages with two formally distinct series of adjectival possessives: a weak and a strong one (the dialects of Sicily and several others, such as Southern Calabrese but also Tuscan and Paduan).
 b. Languages which use the same adjectival possessives in both positions (Italian).
 c. Languages with strong adjectival possessives only (e.g. Campano, Teramano, Salentino).³³

Remarkably, no language seems to display the fourth possible case (weak possessive adjectives only).³⁴

Giorgi and Longobardi (1991) did not provide any insight into whether the difference between the dialects with a reduced weak form and (Standard) Italian is a matter of phonology and inflection or a syntactic difference. However, the dialect of Ragusa, for instance, provides evidence for a distributional distinction: prenominal possessives always precede numerals (14).

- (14) a. *i/sti mo tri libbra*
 the-P/ this-P my three books
 ‘the/these three books of mine’
 b. **i/sti tri mo libbra*

In Italian, instead, prenominal possessives may occur after an emphasized numeral and in a high register, also after several types of adjectives:

- (15) a. *i tre miei studenti migliori*
 the-M.P three my-M.P student-M.P best-M.P
 ‘my three best students’
 b. *i tre migliori miei studenti*

The prenominal adjectival possessive visible in the dialects of Sicily and in Southern Calabrese seems then to be subject to a stricter condition that can be summarized as follows: it must be proclitic on the first element following D,

33. In these dialects, adjectival possessives are rather like the Spanish ones; however, unlike Spanish, which displays weak (definiteness-checking) possessives along with the strong series, these dialects lack any prenominal (weak) alternative.

34. This may be due to the fact that all languages try to have a possessive series available at least in one strong position (i.e. with null head nouns). In Romance, this goal can never be achieved through a definiteness-checking strong possessive (like English *mine*), not even in languages with definiteness-checking weak ones (e.g. French or Spanish) perhaps for principled reasons connected to the syntax of bare nouns and proper names (Longobardi 1996, 54).

making it a syntactic second-position clitic.³⁵ Then, improving on previous treatments, we can conclude that there are three types of possessives: *strong*, *weak* and (second-position) syntactic *clitics*, the latter two occurring both, under slightly different conditions, in the prenominal position.

One may wonder whether the existence of the weak series in prenominal position in Italian, against its uniform absence in all SIDs, can be related to the extent of prenominal placement of adjectives, i.e. to parameter NM1 above.³⁶ An argument in this direction is that no Romance variety has been pointed out with pre- and post-nominal possessives (except for the ‘clitic’ series of Sicily and Southern Calabrese) and only postnominal adjectives. A language like Spanish apparently exhibits the complementary intermediate case, i.e. only postnominal adjectival possessives and both pre- and post-nominal adjectives, but this could be plausibly related to the fact that Spanish has prenominal definiteness-checking possessives, hence with the distribution of determiners. Indeed, notice that there seems not to be a language exhibiting both weak and clitic possessive series either. All these facts amount to the tentative generalization that there are no languages with more than one series of possessives occurring prenominally (the same could possibly be true

35. The cases of fn. 28 (enclitic, uninflected possessives used with kinship nouns and few others which appear to have raised to D) in Salentino, Teramano, Campano and Northern Calabrese could be unified with the more general second-position clitics of Sicilian and many other dialects, if we supposed that the difference between the two types is that in Sicilian they are phonologically proclitic on the noun (or its extended projection) while in the other case they are phonologically enclitic on the noun: this would let them to occur exactly after the few nouns allowed in each language to raise to D (Longobardi 1996). This would make the further (possibly incorrect) prediction that they should be able to encliticize to proper names, and would raise the (correct) expectation of some geolinguistically complementary distribution between the two subtypes. We will not pursue this tentative line in the parametrization for this article.

36. The pre-/post-nominal alternation of the only series of possessives of (Standard) Italian is reminiscent of the pattern of adjectives also because postnominal possessives, at least in definite phrases, exhibit an obligatory restrictive interpretation exactly as postnominal adjectives which may also occur in prenominal position:

- | | | |
|--------|---|------------------------------------|
| (i) a. | <i>il mio libro / il nuovo libro</i> | |
| | the my book / the new book | (both not necessarily restrictive) |
| | b. <i>il libro mio / il libro nuovo</i> | (both necessarily restrictive) |

In all the SIDs, instead, where (non-clitic) possessives and adjectives are only postnominal, the reading need not be restrictive for either of them, exactly as is the case for Italian adjectives which may only occur postnominally (‘la macchina tedesca di Gianni’). Whatever mechanism assigns restrictive readings to adjectives should then probably extend to possessives, under any parametrization.

even postnominally), even if the two or three series could be distinguished syntactically by being weak, clitic or definiteness-checking.³⁷

These patterns may instantiate a particular case of Anti-Synonymy, a potential third-factor principle that has been proposed to apply to other functional elements of the lexicon as well (Keenan 2002, Longobardi 2014). The relevant intuition, restricted here to functional words, can be summed up as follows:

- (16) *Anti-Synonymy*. A language has a single morpho-phonological matrix for each functional meaning.

In the case of different possessive series, Anti-Synonymy can be relaxed as long as the series are morpho-phonological alternatives in the two salient contexts for possessives, the strong and weak position. But if a language has both a prenominal and a postnominal position for possessives and two series of possessives, it suggests that in either position only one series may occur.

Thus, Anti-Synonymy forces Spanish, which has pre- and post-nominal adjectives like Italian, and could thus be expected to also have pre- and postnominal adjectival possessives, to drop them from the prenominal position: in that position, they are the succumbing competitor of definiteness-checking possessives, which are inevitably prenominal (they need to overtly check the D position) and are absent in Italo-Romance as a consequence of an independent parameter (PDC above). This gives rise to the characteristic pattern with a weaker definiteness checking possessive and a stronger adjectival series occurring only postnominally (*mi libro*, *el libro mío*, *un/este libro mío*..., but **el mío/mi libro*, **un/este mío/mi libro*...: Giorgi and Longobardi 1991, Chapter 3, Appendix), distinguishing Spanish from Italian and also from SIDs, without further parametric stipulation.³⁸ In a

37. The generalization could also be supported by the diachronic reduction of contexts acceptable for French adjectival *mien* in correspondence of the rise of *mon* in the past three centuries (Posner 1990) and, conversely, by the residual status of Catalan *mon* in the face of the generalized adjectival *meu* (cf. Fabra 1956, Badia i Margarit 1962, Hualde 1992).

38. French and European Portuguese are likely to instantiate a parametric pattern very similar to the Spanish one at the level of abstraction relevant for the sample of parameters of this article. The actual manifestation of the same system of values is slightly different, owing to some additional idiosyncrasies: in modern French, the adjectival possessive series of *mien* is reduced not just to strong positions, as predicted by Anti-Synonymy with *mon*, but just to the 'strongest' possible position of all, that occurring with an empty noun, and is completely natural only when introduced by a definite article. Therefore, in Table 7, French was assigned (–) for parameter APO, since here we interpret this parameter as applying to items occurring with the relevant property consistently across all types of DPs. In modern European Portuguese, the prenominal position of possessives is also only natural with a definite determiner, and such forms as *o meu (libro)* will be tentatively analyzed here as representing bimorphemic definiteness-checking

similar way, Anti-Synonymy may apply to determine that, among adjectival possessives, morphologically reduced forms in clitic position (as e.g. in the dialects of Sicily) and full pronominal forms (as in Italian) do not occur in the same language.

If so, then, the differences in possessives between Italian and all SIDs can in principle be reduced to parameter NM1, and those between Sicilian/Southern Calabrese and the other SIDs in our sample to a specific parameter allowing syntactic cliticization to the post-D position or not.

This new systematization of the observed variation amounts then to the single parameter in Table 8 (\pm Wackernagel adjectival possessives, WAP), only active in +APO languages, as shown in the first implicational condition (+APO).³⁹ The effect of Anti-Synonymy in blocking the possibility of a language with the same clitic adjectival possessives as in Sicilian or Southern Calabrese and definiteness-checking possessives as in Spanish or French is recorded in the other implicational condition (\neg +PDC), stating that the parameter is only relevant for languages that are not already +PDC.

Table 8. Wackernagel adjectival possessives

	Parameter	Implications	Sicily, SCa	NCa, Sal, Cam, Ter, It
WAP	\pm Wackernagel ⁴⁰ <i>adj. poss.</i>	+APO, \neg +PDC	+	–

As shown in Table 8, the distribution of the values identifies a geographically uniform area in Southern Italy, that contains the dialects of Sicily and Southern

possessives (like French *mon* and Spanish *mi*), while *meu* in *um libro meu* as adjectival possessives like Spanish *mío*.

39. It remains to be seen if the same parametrization should not extend to definiteness-checking possessives. Definiteness-checking possessives, which have the distribution of determiners, are supposedly necessarily weak in Romance, indirectly owing to the parametric status of D in these languages; hence, the problem does not arise. However, in Germanic or in Celtic, in principle, they could be strong only (i.e. in a language with the equivalent of English *mine*, but without *my*), but in fact this possibility does not appear to be attested, thus the current parametrization has been restricted to adjectival possessives.

40. It is unclear if the position of high adjectival possessives can be identified with/related to that termed GenS present in many languages; the hypothesis that their postnominal counterparts are in the GenO position (Silvestri 2013, Guardiano 2014) faces some difficulty in light of the impossibility of any possessive in the postnominal position available for genitives in German, even in the register allowing an exceptional adjectival behavior with demonstratives.

Calabrese, as opposed to the other SIDs. In this particular domain, the unity of the ExtSDs is thus broken.⁴¹

The empirical manifestations of the parameter are listed in Table 9.

Table 9. Empirical manifestations for the parameter WAP

Parameter	Manifestation(s)	(+)
\pm Wackernagel	<i>Initial condition(s)</i>	
<i>adjectival possessives</i>	a. the language has adjectival possessives b. the language does not have definiteness-checking possessives	
	<i>Manifestation(s)</i>	
	The language has two series of adjectival possessives:	YES
	a. one morphologically <i>weaker</i> (i.e. not inflected for gender and for number), preceding the noun, structured adjectives and numerals in the linear order ([Determiner Poss (Num) N]),	
	b. one fully inflected and postnominal.	
	<i>'Elsewhere' condition(s)</i>	
	The manifestations do not apply if the language has both prenominal and postnominal adjectival possessives, and there are prenominal possessives which are not immediately adjacent to a determiner.	

In this way, most of the variation visible in (non-Balkan)⁴² Romance possessives can be reduced to the combination of the values of two parameters PDC and APO along with the effect of the principle of Anti-Synonymy.

The unification of the variable distributional properties of adjectives and possessives under the single parameter NM1 requires a reconceptualization of the latter. Such a unification makes the empirical prediction that +NM1 languages would also exhibit informationally unmarked postnominal adjectival possessives, and that -NM1 languages would rather feature prenominal (non-Wackernagel) adjectival possessives. This prediction is met, for instance, in the dialects of Sicily (+NM1), where adjectival possessives are expectedly found postnominally and are informationally unmarked. Correspondingly, in Italian (a -NM1 language), the canonical position for possessives is, again expectedly, above the target of N-raising.⁴³ Yet, it is not the case that the difference between Italian and the dialects of Sicily may just amount to the higher scope of the raising of the noun alone

41. For a more detailed analysis of the geographical and historical roots of the special relation of Salentino to the UpSDs, cf. Guardiano et al. (2016).

42. Romanian “definite matching” possessives are excluded from the generalizations made in this section.

43. Additionally, adjectival possessives can be postnominal as long as they are informationally marked. Keeping the assumption that N-raising always targets the same functional head in a

above the aforementioned canonical position of possessives. Indeed, the position of prenominal possessives in Italian is normally before cardinal and ordinal numerals, but these items still precede the noun, that has supposedly raised over possessives in Sicily:

- (17) a. *i miei primi tre figli* [Italian]
 the-M.P my-M.P first-M.P three children-M.P
 'my first three children'
- b. i. *i primi tri figghi mjèi* [Ragusa]
 the-P first-P three children-P my-P
 'my first three children'
- ii. **i figghi primi tri mjèi*
- iii. **i primi figghi tri mjèi*

The position of prenominal possessives in Italian (17a) can be derived as in (18a), where the possessive raises from its thematic position to a pre-numeral (and post-D) one, while the noun does not raise over M1 As. A derivation like (18b), however, would wrongly predict a *N Poss Ord Card* order for the dialects of Sicily.

- (18) a. Italian:
 D Poss Ord Card ...M1... N ...Poss
 ↑ _____|
- b. SIDs:
 *D N Poss Ord Card ...M1... N ... Poss
 ↑ _____|
 ↑ _____|

One simple way to obtain the linear order in (17b.i) is remnant-movement of a wider constituent (rather than simple head-raising): the noun moves over the high possessive along with its higher projections hosting ordinal and cardinal numerals, as in (19).

- (19) SIDs:
 D [Ord Card ... NM1...Poss] Poss [Ord Card ... N M1... Poss]
 _____|
 ↑ _____|
 ↑ _____|

Cast in these terms, parameter NM1 is intrinsically different from the two other N-raising parameters mentioned above (i.e. NOA and NM2), in terms of the type and the size of the moving element: only NM1 predicts movement of a higher projection containing N. This suggestion, or any alternative hypothesis, may imply

certain language, it must be therefore assumed that in Italian two positions for possessives are available: the 'canonical' position and one below the target of N-raising.

ramifications that go beyond the scope of the present work, for which it is sufficient to have shown some possible empirical advantages of parametrically unifying the variable pattern of adjectival possessives and regular adjectives.

This revision would also expand the set of empirical data manifesting parameter NM1, as shown in Table 10.

Table 10. Additional empirical manifestation for the parameter N over M1 As

Parameter	Manifestation(s)	(+)
$\pm N$ over M1 As	In argument position: e. adjectival possessives occur postnominally, as informationally unmarked	YES

3.4 Kinship terms

The interpretation that (certain types of) kinship terms receive in (some) Romance languages indicates that they contain an unpronounced possessive. (Standard) Italian, for instance, allows a 3rd person reading of such an implicit argument. In (20a) *la madre* means ‘his mother. Notice that the implicit argument in Italian has the same readings as *suo/sua* (i.e. both bound and unbound interpretation) rather than *proprio/propria* (bound interpretation only), as shown in (20b).

- (20) a. *Gianni deve accompagnare la madre all' aeroporto*
Gianni must drive the mother to.the airport
‘John must drive his mother to the airport.’
- b. i. *ho avvertito Gianni che Maria ha accompagnato la madre*
have.1 informed Gianni that Maria has driven the mother
all' aeroporto
to.the airport
‘I informed Gianni that Mary has driven his/her mother to the airport’
- ii. *ho avvertito Gianni che Maria ha accompagnato sua*
have.1 informed Gianni that Maria has driven his/her
madre all' aeroporto
mother to.the airport
‘I informed Gianni that Mary has driven his/her mother to the airport’
- iii. *ho avvertito Gianni che Maria ha accompagnato la propria*
have.1 informed Gianni that Maria has driven the own
madre all' aeroporto
mother to.the airport
‘I informed Gianni that Maria has driven her mother to the airport’

A similar interpretation of the implicit argument as 3rd person is observed in Northern Calabrese (21) and Campano.

- (21) *haj ditt a Gjannin ka Màrij ha purtwàt a mughjèra alla stinzjuna*
 have.1 told to Gianni that Mario has driven the wife at.the station
 ‘I told Gianni_i that Mario_j has driven his_{i,j} wife to the station.’

In the dialects of Sicily, Salentino and Southern Calabrese, instead, implicit possessors are ungrammatical:⁴⁴ in sequences like (22) an overt possessive is required.

- (22) *Gjovànni ha purtàri *â mughjèri / a sso muggjèri â*
 Gianni has drive DOM.the wife / DOM his wife to.the
stazzjòni [Ragusa]
 station
 ‘John has to drive his wife to the station.’

This same pattern is well-attested outside of Italo-Romance, e.g. in Spanish (23):

- (23) *Juan debe acompañar *a la mujer/ a su mujer al aeropuerto*⁴⁵
 Juan must drive DOM the wife/ DOM his wife to.the airport
 ‘John has to drive his wife to the airport.’

These data suggest that kinship terms come with an implicit argument in some Romance languages (e.g. Italian, Campano, Northern Calabrese) but not in others (e.g. the ExtSDs or Spanish). Such a crosslinguistic distinction seems not to be reducible to any other parametric difference. Thus, very descriptively, in order to capture it, we postulate parameter *±obligatory possessive with kinship N* (OPK, Table 11): the value (+) of this parameter predicts the incompatibility of kinship

44. The nouns for ‘mum’ and ‘dad’ (and ‘grandma’/‘grandpa’ or even sometimes ‘uncle’/‘aunt’) are exceptional in at least two respects: (a) in most of the languages considered (but crucially not in Sicilian), they are likely to raise to D (Longobardi 1996) and need an implicit argument understood as 1st/2nd person (sentence in (i) below); (b) in Sicilian and other dialects they have an unpronounced possessive understood as 1st/2nd person (sentence in (ii) below).

(i) *devo accompagnare mamma alla stazione* [(Standard) Italian]
 I.must drive mum to.the station
 ‘I have to drive (my/our) mum to the station.’

(ii) *ha purtàri â mamà â stazzjoni* [Ragusa]
 I.must drive dom.the mum to.the station
 ‘I have to drive (my/our) mum to the station.’

45. Some speakers do not find this sentence entirely ungrammatical although they agree it is more restricted: both *Juan* and *his wife* must have been already introduced in the discourse, for it to be usable.

nouns with implicit arguments (as in the ExtSDs and Spanish) and, as a consequence, the requirement of a visible possessive (Table 12); the value (–) implies the possibility for a language to license an (optional) 3rd person implicit argument. Notice that the distribution of the two values across our sample of SIDs reflects the traditional partition between ExtSDs and UpSDs: this is the only case in our sample that shows such a neat separation.

Table 11. Obligatory possessive with kinship nouns

	Sicily, Sal, SCa, Sp	It, Ter, Cam, NCa, Rm, ...
OPK ± <i>obligatory possessive with kinship nouns</i>	+	–

The empirical manifestation(s) of the value (+) of this parameter are listed in Table 12.

Table 12. Obligatory possessive with kinship nouns: manifestations

Parameter	Manifestation(s)	(+)
± <i>obligatory possessive with kinship nouns</i>	A visible (3rd person) possessive is used with singular kinship nouns (other than ‘mum’ and ‘dad’) in order to obtain a (3rd person) possessor reading.	YES

3.5 Summary

The distribution of the values of the parameters described in this section across the dialects of our sample defines different degrees of variability and largely reflects their traditional classification (which is fundamentally based on non-syntactic evidence). With the exception of FVP, whose distribution, as mentioned above, seems not to correlate with any sociolinguistic or geographical variable, the other parameters define syntactic isoglosses which are compatible with well-established dialectological patterns. For instance, one parameter (i.e. PGO) singles out one individual language (Northern Calabrese) from the broader Southern Italy unity, thus predicting its position as an isolate; similarly, parameter EAF differentiates Southern Calabrese from its closest relatives (i.e. the other dialects of the broader Sicilian area); parameter NM1, on the contrary, establishes a Southern Italo-Romance unity that excludes (Standard) Italian; two other parameters, finally, seem to display a geographically-driven distribution (i.e. WAP, that separates the extended Sicilian area from the rest, and FNN, that singles out the two “Northern” dialects of the sample, i.e. Campano and Teramano). In view of this distribution, a taxonomy such as the one in Figure 1 is in fact largely predictable.

4. Conclusions

Building on a previous PCM-based study (Guardiano et al. 2016), this paper confirms that the PCM is able to compare languages even at the level of microvariation. In order to do so, the system must be enhanced with additional parameters able to capture microscopic aspects of variation. Such parameters retrieve independently known taxonomic results, suggest some novel plausible historical details (e.g. the effects of prolonged geographical isolation vs. ancient road connection to Rome and Naples on different dialects of the same branch, suggested in Guardiano et al. 2016), and inspire improved analyses of other Romance varieties.

Furthermore, considering microvariation within the PCM framework refines our understanding of the complex ways parameters are mapped onto manifestations and ultimately onto empirical data. In the paper, we provided examples of such interaction through the description of some aspects of variability in the nominal domain.

Acknowledgements

This paper is largely based on joint work with Giuseppe Longobardi, who has significantly contributed in particular to Sections 2, 3.1, 3.3.

This work was partly supported by the ERC Advanced Grant 295733 LanGeLin (Principal Investigator: G. Longobardi). The authors are grateful to Birgit Alber, Andrea Ceolin, Richard Kayne, Edward Keenan, Hilda Koopman, Iara Mantenuto, Giuseppina Silvestri and Melita Stavrou for helpful discussions on various parts of the analysis and the data. We also thank two anonymous referees.

References

- Alexiadou, Artemis. 2014. *Multiple determiners and the structure of DPs*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.211>
- Alexiadou, Artemis, Liliane Haegeman, and Melita Stavrou. 2007. *The Noun Phrase in the Generative Perspective*. Berlin: De Gruyter. <https://doi.org/10.1515/9783110207491>
- Arteaga, Deborah. 1995. "On Old French Genitive Constructions." In *Contemporary Research in Romance Linguistics*, ed. by Jon Amastae, Grant Goodall, Mario Montalbetti, and Marianne Phinney, 79–90. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.123.08art>
- Badia i Margarit, Antoni M. 1962. *Gramàtica catalana*. Madrid: Gredos.
- Bernstein, Judy. 1991. "DPs in French and Walloon. Evidence for Parametric Variation in Nominal Head Movement." *Probus* 3.2: 101–126. <https://doi.org/10.1515/prbs.1991.3.2.101>
- Cardinaletti, Anna. 1998. "On the Deficient/Strong Opposition in Possessive Systems". In *Possessors, Predicates, and Movement in the Determiner Phrase*, ed. by Artemis Alexiadou, and Chris Wilder, 17–53. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.22.03car>

- Cinque, Guglielmo. 2010. *The Syntax of adjectives. A Comparative Study*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9780262014168.001.0001>
- Crisma, Paola. 1991. "Functional Categories Inside the Noun Phrase. A Study on the Distribution of Nominal Modifiers." M.A. Thesis, Università di Venezia.
- Crisma, Paola. 1996. "On the Configurational Nature of Adjectival Modification." In *Grammatical Theory and the Romance Languages*, ed. by Karen Zagona, 71–93. Amsterdam/Philadelphia: John Benjamins. <https://doi.org/10.1075/cilt.133.06cri>
- Crisma, Paola, and Chiara Gianollo. 2006. "Where did Romance N-raising come from? A parallel study of parameter resetting in Latin and English." In *Romance Languages and Linguistic Theory 2004: Selected papers from 'Going Romance'*, ed. by Jenny Doetjes and Paz González, 71–93. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.278.05cri>
- D'Alessandro, Roberta, and Laura Migliori. 2013. "Sui possessivi enclitici nei dialetti italiani meridionali." Ms., Leiden University.
- Delfitto, Denis, and Paola Paradisi. 2009. "Prepositionless Genitive and N+N Compounding in (Old) French and Italian." In *Romance Languages and Linguistic Theory 2006*, ed. by Danièle Torck, and W. Leo Wetzels, 53–72. Amsterdam: John Benjamins.
- Fabra, Pompeu. 1956. *Gramàtica catalana*. Barcelona: Teide.
- Foulet, Lucien. 1919. *Petite syntaxe de l'ancien français*, Paris: Honoré Champion.
- Gianollo, Chiara. 2009. "Prepositional Genitives in Romance and the Issue of Parallel Development." Talk given at XI Diachronic Generative Syntax Conference, Universidade Estadual de Campinas (DiGS), July 22–24, 2009.
- Giorgi, Alessandra, and Giuseppe Longobardi. 1991. *The Syntax of Noun Phrases: Configuration, Parameters and Empty Categories*. Cambridge: Cambridge University Press.
- Guardiano, Cristina. 2014. "Fenomeni di contatto sintattico in Italia meridionale? Alcune note comparative." In *Quaderni di lavoro ASIt* 18: 73–102.
- Guardiano, Cristina, and Giuseppe Longobardi. 2005. "Parametric Comparison and Language Taxonomy." In *Grammaticalization and Parametric Variation*, ed. by Monserrat Batllori, Maria-Lluïsa Hernanz, Carmen Picallo, and Francesc Roca, 149–174. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199272129.003.0010>
- Guardiano, Cristina, and Giuseppe Longobardi. 2017. "Parameter Theory and Parametric Comparison." In *The Oxford Handbook of Universal Grammar*, ed. by Ian Roberts, 377–398. Oxford: Oxford University Press.
- Guardiano, Cristina, Dimitris Michelioudakis, Andrea Ceolin, Monica Alexandrina Irimia, Giuseppe Longobardi, Nina Radkevich, Giuseppina Silvestri, and Ioanna Sitaridou. 2016. "South by Southeast: A Syntactic Approach to Romance and Greek Microvariation." *L'Italia Dialettale* LXXVII, 96–166.
- Guardiano, Cristina, and Melita Stavrou. 2014. "Greek and Romance in Southern Italy: History and Contact in Nominal Structures." *L'Italia dialettale* LXXV: 121–147.
- Guardiano, Cristina, and Melita Stavrou. In preparation. "Adjective-Noun Combination in Romance and Greek of Southern Italy. Polydefiniteness Revisited."
- Hualde, José Ignacio. 1992. *Catalan*. Oxon: Routledge.
- Irimia, Monica Alexandrina, Giuseppe Longobardi, Dimitris Michelioudakis, and Nina Radkevich. in prep. "Romanian Genitives and Universal Grammar."
- Keenan, Edward. 2002. "Explaining the creation of reflexive pronouns in English." In *Studies in the History of the English Language*, ed. by Donka Minkova, and Richard Stockwell, 325–354. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110197143.3.325>

- Laenzlinger, Christopher. 2005. "French Adjective Ordering: Perspectives on DP-internal Movement Types." *Lingua* 115: 645–689. <https://doi.org/10.1016/j.lingua.2003.11.003>
- Lausberg, Heinrich. 1939. *Die mundarten Suedlukaniens*. Halle: Niemeyer.
- Ledgeway, Adam. 2016. "The dialects of southern Italy." In *The Oxford guide to the Romance languages*, ed. by Adam Ledgeway, and Martin Maiden, 246–269. Oxford: Oxford University Press.
- Longobardi, Giuseppe. 1996. *The syntax of N-raising: a minimalist theory*. OTS Working Papers, Research Institute for Language and Speech, University of Utrecht.
- Longobardi, Giuseppe. 2003. "Methods in Parametric Linguistics and Cognitive History." *Linguistic Variation Yearbook* 3: 101–138. <https://doi.org/10.1075/livy.3.06lon>
- Longobardi, Giuseppe. 2014. "Theory and experiment in Parametric Minimalism: the case of Romance negation." In *Language Description Informed by Theory*, ed. by Rob Pensalfini, Myfany Turpin, and Diana Guillemin, 217–261. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.147.10lon>
- Longobardi, Giuseppe, Cristina Guardiano, Giuseppina Silvestri, Alessio Boattini, and Andrea Ceolin. 2013. "Toward a syntactic phylogeny of modern Indo-European languages." *Journal of Historical Linguistics* 3: 122–152. <https://doi.org/10.1075/jhl.3.1.07lon>
- Longobardi, Giuseppe, and Cristina Guardiano. 2009. "Evidence for Syntax as a Signal of Historical Relatedness." *Lingua* 119 (11): 1679–1706. <https://doi.org/10.1016/j.lingua.2008.09.012>
- Longobardi, Giuseppe, and Cristina Guardiano. 2017. "Phylogenetic Reconstruction in Syntax: The Parametric Comparison Method." In *The Cambridge Handbook of Historical Syntax*, ed. by Adam Ledgeway, and Ian Roberts, 377–398. Cambridge: Cambridge University Press.
- Longobardi, Giuseppe, Monica Alexandrina Irimia, Dimitris Michelioudakis, and Nina Radkevich. 2016. "On Genitive and Other Checked Cases." Paper presented at *Nominals 2016*. Recife, Brazil.
- Longobardi, Giuseppe, and Giuseppina Silvestri. 2013. "The structure of noun phrases." In *The Bloomsbury Companion to Syntax*, ed. by Silvia Luraghi, and Claudia Parodi, 88–117. London: Bloomsbury.
- Loporcaro, Michele. 2009. *Profilo linguistico dei dialetti italiani*. Bari: Laterza.
- Maiden, Martin, and Mair Parry. 1997. *The Dialects of Italy*. London: Routledge.
- Mantenuto, Iara. 2015. "A syntactic analysis of epenthetic [-a] in Teramano." Ms., UCLA.
- Manzini, Rita, and Leonardo Savoia. 2007. *I dialetti italiani e Romanci, Morfosintassi generativa*. Alessandria: Edizioni dell'Orso.
- Martino, Paolo. 1991. *L' "area Lausberg" fra isolamento e arcaicità*. Roma: Dipartimento di Studi Glottoantropologici, Università di Roma 'La Sapienza'.
- Pellegrini, Giovan Battista. 1977. *Carta dei dialetti d'Italia*. Pisa: Pacini.
- Posner, Rebecca. 1990. "Linguistics and philology: parametric changes and Romance possessives." In *Historical linguistics and philology*, ed. by Jazek Fisiak, 337–352. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110847260.337>
- Roberts, Ian. 2012. "On the Nature of Syntactic Parameters: A Programme for Comparative Research." In *Parameter Theory and Linguistic Change*, ed. by Charlotte Galves, Sonia Cyrino, Ruth Lopes, Filomena Sandalo, and Juanito Avelar, 319–334. Oxford: Oxford University Press.
- Rohlf, Gerhard. 1966. *Grammatica storica della lingua italiana e dei suoi dialetti*. Torino: Einaudi.
- Rouveret, Alain. 1994. *Syntaxe du gallois. Principes généraux et typologie*. Paris: CNRS Editions.
- Salvi, Giampaolo, and Lorenzo Renzi. 2010. *Grammatica dell'Italiano antico*. Bologna: Il Mulino.

- Scott, Gary-John. 2002. "Stacked Adjectival Modification and the Structure of Nominal Phrases." In *Functional structure in the DP and IP: The cartography of syntactic structures volume I*, ed. by Guglielmo Cinque, 91–120. Oxford: Oxford University Press.
- Silvestri, Giuseppina. 2013. *The Nature of Genitive Case*. Ph.D. dissertation, Università di Pisa.
- Sitaridou, Ioanna. 2017. "Objects." In *Manual of Romance Morphosyntax and Syntax*, ed. by Andreas Dufter, and Elisabeth Stark, 89–153. Berlin: De Gruyter Mouton.
- Sproat, Richard, and Chilin Shih. 1991. "The Crosslinguistic Distribution of Adjective Ordering Restrictions." In *Interdisciplinary approaches to language*, ed. by Carol Georgopoulos, and Roberta Ishihara, 565–593. Dordrecht: Kluwer.
https://doi.org/10.1007/978-94-011-3818-5_30
- van Rijn, Marlou. 2016. "The Grammaticalization of Possessive Person Marking: A Typological Approach." *Transactions of the Philological Society* 114 (2): 233–276.

Appendix



Map 1. The Italian dialects in this study.

Table A. Number on N

		RGS, AdS, MsS, Sal, SCa, NCa, It, Sp, Ptg, Rm	Cam, Ter, Fr
FNN	\pm Number on N	+	-

Table B. Manifestations of Number on N (either of the two conditions is sufficient in order for the value + to be set. The value - can be set by default)

\pm Number on N	(+)
<i>Initial condition</i>	
The language has obligatory number morphology on some grammatical element.	
<i>Manifestations</i>	
a. Number morphology is robustly visible on all classes of nouns.	YES
b. At least some types of common nouns (e.g. plural nouns with indefinite reading, or mass nouns with indefinite readings, etc.) can be bare (i.e. determinerless) in argument position when occurring with no adjectival modifiers	YES

Table C. Variable person in D

		RGS, AdS, SCa, NCa	MsS, Sal, Cam, Ter, It
FVP	\pm variable Person in D	+	-

Table D. Empirical manifestations of Variable Person in D (either of the two conditions is sufficient in order for the value + to be set. The value - can be set by default)

\pm variable Person in D	(+)
<i>Manifestations</i>	
a. A DP non-overtly marked as 1st/2nd person in subject position can control non-3rd person verbal agreement.	YES
b. A DP non-overtly marked as 1st/2nd person in object position can be doubled by a 1st/2nd person clitic.	YES

Morphological doublets in Brazilian Portuguese *wh*-constructions

Mary A. Kato
State University of Campinas

This chapter contains an overview of diachronic and synchronic studies on Brazilian Portuguese *wh*-questions, and discusses the intriguing problem of apparent syntactic doublets, or optionality, in this domain. However, a careful examination of the diachronic order in which these patterns appear, and the variation exhibited in contemporary written and spoken corpora lead us to conclude that the co-occurring patterns do not constitute syntactic “doublets.”

Keywords: *wh*-questions, Brazilian Portuguese, doublets, optionality, diachrony, *wh*-movement

1. Introduction

Contemporary Brazilian Portuguese (BP) exhibits variation in *wh*-constructions as in Example (1). At first sight, this faces a problem concerning the *wh*-parameter, according to which languages can be of two types: with or without *wh*-movement (Huang 1982).

- (1) a. *Você votou em quem em 1989?* BP
you voted for who in 1989?
b. *Em quem você votou em 1989?* BP
in whom you voted in 1989
'Who did you vote for in 1989?'

Moreover, the VS order in *wh*-constructions is often optional, as can be seen in Example (2).¹

1. But VS order with transitive verbs is ruled out in BP, while it is the norm with all types of verbs in EP (cf. Ambar 2003, Kato and Tarallo 2003, a.o.). But EP and BP behave in the same way, with SV order, when we have D-linked *wh*-questions (Ambar 1988).

- (2) a. *Afinal, quanto vale a vida?* BP
 after-all, how-much is-worth the life
- b. *Afinal, quanto a vida vale?* BP
 after-all how-much the life is-worth
 ‘After all, how much is life worth?’

What we have in BP *wh*-constructions is a three-way variation as in (3), which should not be allowed considering the “Blocking Effect” (Aronoff 1976), a principle that excludes morphological or syntactic “doublets” or, in this case, “triplets.”

According to Kroch (1994: 181) “doublets are always reflections of unstable competition between mutually exclusive grammatical options,” which normally result from grammatical change.

- (3) a. *Onde dormem as crianças?* [*wh*-VS order] BP EP
 where sleep the children
- b. *Onde as crianças dormem?* [*wh*-SV order] BP *EP
 where the children sleep
- c. *As crianças dormem onde?* [*wh*-in-situ] BP %EP²
 the children sleep where
 ‘Where do the children sleep?’

Embedded questions, on the other hand, exhibit a uniform pattern, with fronted-*wh* constituents, and no subject-verb inversion in European Portuguese (EP) or BP:

- (4) *A Ana está perguntando onde as crianças dormem.* BP EP
 the Ana is asking where the children sleep
 ‘Ana is asking where the children sleep.’

My aim in this paper is to show that the three variants in (3) do constitute reflexes of change, but what I want to claim is that they do not constitute a “triplet” and can be analyzed as possible stylistic variants within the same grammar.

The chapter is organized as follows. Section 2 contains an overview of the diachronic data in Portuguese. Section 3 contains the analysis of the data. The last section contains the conclusions.

2. European Portuguese (EP) has a much more restricted (%) occurrence of *wh*-in-situ, part of which with infinitives (Lopes-Rossi 1996). Ambar (2003) corroborates this fact. We will leave aside *wh*-in-situ in Modern EP in this paper.

2. The diachronic development of *wh*-constructions in BP

2.1 The V2 period

Until the end of 18th century, or the Classic period, *wh*-questions in Portuguese³ had the common V2 (*Wh-VS*) pattern of Romance in general,⁴ a pattern that continues to subsist in Modern European Portuguese (EP), but only survives as a residue in BP. Moreover, the VS order that we find in the 20th century is only of the type in (6) with a DP subject (cf. Duarte and Kato 2002), and not a pronominal one.

- (5) a. *Mas que me quer ele a mi?* (16th) Classic Portuguese, CIP
 but what me want he of me
 ‘But what does he want from me?’
- b. *Que fez vossa mercê?* (18th) CIP
 what did your highness
 ‘What did your highness do?’
- c. *Que te disse ele?* (19th) EP
 what you-accus said he
 ‘What did he say?’
- d. *Mas que tenho eu com isso?* (19th) BP
 but what have I with it
 ‘But what do I have with it?’
- (6) a. *Onde andar  a Neiva?* (19th–20th) EP BP
 where will-walk the Neiva
- b. *Que lhe disse o Hoorato?* (19th–20th) EP BP
 what you-cl said the Honorato
 ‘What did Honorato tell you?’

3. We only distinguish between European Portuguese and Brazilian Portuguese from the 19th century on.

4. It also had a very restricted occurrence of pseudo-cleft *wh*-questions, which we will not include in our analysis.

- (i) *QUE he o que dizes, irm a?* (14th)
 what is what say.2SG sister
 ‘What do you say, sister?’
- (ii) *Que   o que quer?* (17th)
 what is what want.2SG
 ‘What is it that you want?’

2.2 The inverse [é que] cleft construction of modern Portuguese

Since the second half of the 17th century, a *wh*-construction started to appear from a different numeration, in both BP and European Portuguese (EP), namely a cleft *wh*-sentence (Example (7a)), with the copula having the privileged second position and the thematic verb exhibiting the SV order of its embedded clause (Duarte 1992, Duarte and Kato 2002). The grammar that produces the inverse cleft *wh*-question is the same as the V2 grammar, with *wh*-movement, and the copula in second position.⁵ I consider this a kind of grammaticalization as in the V2 grammar any thematic verb could be raised to C, and in the cleft case this possibility is reduced to the copula, a functional verb.

From the 19th century on, the copula ceases to present *consecutio temporum*, exhibiting the invariable form *é* (Example (7b)), again a further case of grammaticalization.

- (7) a. *Onde foi que as crianças dormiram?*
 where was that the children slept
 ‘Where was it that the children slept?’ [inverse cleft, BP/EP 18th]
- b. *Onde é que as crianças dormiram?*
 where is that the children slept
 ‘Where was it that the children slept?’ [inverse cleft, BP/EP 20th]

2.3 The *wh-in-situ* construction

While EP has a very restricted occurrence of *wh-in-situ* questions (1.22% in written data), the second half of the 19th century sees, in BP, the appearance of a much larger amount (8.84 in the written corpus) (Kato and Mioto 2005), achieving 42% of occurrences in the spoken Brazilian data of the 20th century (Lopes-Rossi 1996). At this point, we should stress that BP distinguishes two types of *in-situ wh*-constructions: (a) the one with rising intonation (↑) (8a), with echo interpretation, and the one with falling intonation (↓) (8b) a genuine question (cf. Kato 2013a, 2013b).⁶

5. Comparing the cleft sentence derivation with the V2 one, we have:

- (i) [_{CP} onde [_{dormem} [_{TP} as crianças ~~dormem~~ [_{VP} as criança ~~dormem~~ onde]]]]
- (ii) [_{CP} onde [foi [_{TP} foi [_{VP} foi [_{CP} que as crianças dormiram onde]]]]]]

6. With falling intonation, the *wh*-element does not have to be at the sentence final position as in the echo-question.

- (8) a. *Você é de onde?*↑↑
 you are from where
 [Rising intonation, spoken BP]
- b. *Você pensava em que na sua vida?*↓↓
 you thought in what in.the your life
 [Falling intonation, spoken BP]

Although BP and French (FR) can be considered similar types of language, licensing both moved-*wh* and *wh-in-situ*, we have to explain why BP allows (9b and c) while French does not license the parallel patterns (10b) and (10c) (Cheng and Rooryck 2000, Vergnaud and Zubizarreta 2005). Moreover, French has the same intonation for both echo and genuine *in-situ* constructions.

- (9) a. *então você vai dizer o que?*↓↓
 so you go say what
 ‘So what are you going to say?’ [BP, second half, 19th]
- b. *E Maria pensa que João comprou o que?*↓↓
 and Mary thinks that John bought what
 ‘And what does Mary think that John bought?’
 [BP, *in-situ* in complement clause]
- c. *Maria ama o livro que quem escreveu?*↓↓
 Mary loves the book that who wrote
 Lit: ‘Mary loves the book that who wrote?’ [BP, *in-situ* inside an island]
- (10) a. *alors tu vas dire quoi?*↑ FR
- b. **Marie pense que Jean a acheté quoi?*↑ [FR, *in-situ* in complement clause]
- c. **Marie aime le livre que qui a écrit?*↑ [FR, *in-situ* inside an island.]

2.4 The semi-cleft *wh*-construction

A change that follows the introduction of the *wh-in-situ* construction, in the second half of the 19th century, is the cleft type without the copula, the semi-cleft, existent only in BP, and observed to occur mostly in the 20th century spoken language by Lopes-Rossi (1996) and Duarte and Kato (2002):

- (11) a. *O que que você faz?* (20th, spoken BP)
 what that you do
 ‘What do you do?’
- b. *De quem que é esse peixe?* (20th, spoken BP)
 of whom that is this fish
 ‘whose is this fish?’

2.5 The appearance of the pattern *wh*-SV

In the second half of the 19th century, BP sees also the appearance of the non-V2 type of construction, namely the *wh*-SV pattern, a unique type among Romance languages, which occurs both in written and in spoken language.

- (12) a. *Onde ele foi?* (19th, written BP)
 where he went
 ‘Where did he go?’
- b. *Quanto você ganha?* (20th, spoken BP)
 how much you earn
 ‘How much do you earn?’

Brazilians avoid the Reduced Cleft (*wh-que*-SV) in written language while they prefer it in colloquial speech. What Brazilians use instead, in written language, is the *wh*-SV type, prescribed by traditional grammarians as the prototypical Brazilian norm. This point will be treated in detail below in 3.7.

Summarizing the diachrony of *wh*-questions in BP, Lopes-Rossi (1996) draws Figure 1:

OP & CIP	<i>wh</i> VS				
14th–18th					
EP	<i>wh</i> VS	<i>wh é que</i>	% <i>wh-in-situ</i>		
18th–20th		VS/SV			
BP	-----	<i>wh é que</i>	<i>wh-in-situ</i>	<i>wh que</i>	<i>wh</i> SV
18th–20th		SV		SV	

Figure 1. Types of *wh*-questions in Portuguese through time (adapted from Lopes-Rossi 1996)⁷

2.6 A missing cell between the *wh-in-situ* and the *wh-que* SV

Assuming that Focus constructions and *wh*-constructions have parallel derivations due to the fact that they both check their features against the same head Foc, Kato and Ribeiro (2009) noticed a gap between the *in-situ* construction and the semi-cleft pattern.

What they found was that in the V2 period, the cleft type was the inverse type, with the copula in second position, but with the advent of *wh-in-situ* the cleft type

7. OP = Old Portuguese, CIP = Classic Portuguese

also changes from the inverse type to a canonic type of cleft, with the copula no longer in second position.

- (13) a. *E ISSO é que se chama postura, ou posição reta.*
and this is that se called posture or position straight
'And that is what is called posture, or straight position.'
(17th, inverse cleft)
- b. *E QUANDO é que são relativas?*
and when is that are relatives
'And when are they relatives?'
(17th, inverse cleft)
- (14) a. *É a Maria que tá tocando violão.*
is the Mary that is playing guitar
'It is Mary that is playing the guitar.'
(20th, canonic cleft)
- b. *É quem que tá tocando o violão?*
is who that is playing the guitar
'Who is it that is playing the guitar?'
(20th, canonic cleft)⁸

We can modify the Lopes-Rossi's previous Figure, adding the canonic cleft pattern between the *wh-in-situ* and the *wh-que SV* pattern (see Figure 2):

OP & CIP 14th–18th	<i>WhVS</i>					
EP 18th–20th	<i>WhVS</i>	<i>Wh é que VS/SV</i>	<i>%(wh-in-situ)</i>			
BP 18th–20th	<i>% WhVS</i>	<i>Wh é que SV</i>	<i>wh-in-situ</i>	<i>é Wh que</i>	<i>Wh que SV</i>	<i>Wh SV</i>

Figure 2. Types of BP *WH*-questions through time

3. Syntactic analyses

3.1 Thetic and categorial sentences

I assume, with Kuroda (1972), that languages distinguish between Categorical and Thetic sentences. In EP this distinction is expressed through the word order *SV* and *VS* (Martins 1994), a contrast that has been claimed to be lost in BP (Britto

8. The canonic type of *wh*-question was not found in the corpora used, a reason why we find a gap in Lopes-Rossi's Figure 1. Kato (2013b, 2014) found examples in dialogues between a child and a mother. Such examples were also easily accepted by my undergraduate students.

2000), according to which the order SV(X) expresses athetic judgment, while the categorical judgment is codified as a Left Dislocation (LD) structure.

- (15) a. *O gato está a dormir.* SV (categorical) EP
 the cat is to sleep
 b. *Está o gato a dormir.* VS (thetic) EP
 is the cat to sleep
 ‘The cat is sleeping.’
- (16) a. *O gato, ele está dormindo.* LD (categorical) BP
 the cat he is sleeping
 b. *O gato está dormindo.* SV (thetic) BP
 the cat is sleeping
 ‘The cat is sleeping.’

3.2 The oldest type of *wh*-questions in Portuguese: The V2 construction

I also assume that *WH*-questions are thetic sentences and will, therefore, mirror the order VS of declarative thetic sentences in CIP. In EP, the pattern is still part of the vernacular. See (18a)’ for the derivation of a declarative sentence (Kato 1987, Martins 1994). As for a *wh*-question, the derivation in (17b)’ mirrors that of a declarative sentence, concerning the position of the finite verb.

- (17) a. *Está o gato a dormir no jardim.* CIP EP
 is the cat to sleep in the garden
 ‘There is a cat sleeping in the garden.’
 b. *Onde está o gato a dormir?* CIP EP
 where is the cat to sleep
 ‘Where is the cat sleeping?’
- (17)⁹ a. [_{FocP} [_F *está* [_{IP} *está* [_{VP} *o gato está* [_{VP} *está a dormir no jardim*]]]]]
 b. [_{FocP} *onde* [_F *está* [_{TP} *está* [_{VP} *o gato está* [_{VP} *está a dormir onde*]]]]]]]

My analysis assumes, with Uriagereka (1995) and Kato and Raposo (1996), that the finite verb in Western Iberian Romance, moves to F, in the left periphery of the sentence, which also happens in Modern European Portuguese (EP).¹⁰

BP, on the other hand, has lost the V2 type of *wh*-constructions with thematic verbs, and the apparent contemporary V2 type has been analyzed in Kato (1993) as a stylistic inversion in French (Kayne and Pollock 1978), or a right dislocation

9. I will ignore the projection ForceP in both derivations.

10. A different version is given in Barbosa (2001) for whom the verb is in T and the subject *in-situ*, in *Spec, v*.

with the subject null in Duarte and Kato (2002). This explains why the subject is usually a name or a DP.¹¹

- (18) *Onde (ela) está essa mulher?* 20th BP
 where she is this woman 12

(18') [_{FocP} onde [_{TP} [_{TP} (ela) está [_{vP} (ela) ~~está onde~~]] essa mulher]

3.3 The inverse-cleft questions (*WH-é que SV*) as a grammaticalized V2 construction

Until the old and the classic period (the V2 period), the *wh*-element in Portuguese was followed by any thematic or auxiliary verb, a fact that changed in the 18th century as the copula started to have the privileged second position, a process that we can consider a type of grammaticalization.

- (19) a. *Que compraste (tu)?* CIP EP *BP
 what bought you
 'What did you buy?'
 b. *Quando chegaram as crianças?* CIP EP %BP
 when arrived the children
 'When did the children arrive?'
 c. *Onde podem as crianças estar a dormir?* CIP EP *BP
 where can the children be sleeping
- (20) a. *Que foi que tu compraste?* EP BP
 what was that you bought
 'What was it that you bought?'
 b. *Quando é que chegam as crianças?* EP BP
 when is that arrive the children
 'When do the children arrive?'

The order of the thematic verb with the subject no longer matters, as the former occurs inside an embedded sentence. What now moves to F is the copula.

(20') a. [_{FocP} que [_Ffoi [_{TP} foi [_{vP} foi [foi [_{CP} que tu compraste ~~que~~]]]]]]]

11. Actually, as the right dislocation is a type of clitic-doubling phenomenon, the actual derivation should merge the pronoun *ela* with *essa mulher* [(*ela*) *essa mulher*], and *essa mulher* would be the predicate of the "big DP" in Kayne's (2001) sense. Case in this big DP depends on the case of the subject *ela*, a nominative.

12. The presence or absence of the subject pronoun does not make any difference in meaning here.

- b. [_{FocP} quando [_F é [_{TP} é- [vP é- [é- [_{CP} que chegam as crianças quando]]]]]]] quando]]]]]]]

Our prediction is that, while the inverse cleft is the vernacular form in EP, in BP it will survive in written language, being replaced by the shorter semi-cleft form in spoken language.

3.4 From *wh*-movement to a fake “*wh-in-situ*”

In this work, I will assume the cartographic model, namely a hierarchical organization of syntactic constituents, both at the level of clause structures (IP/TP), at the sentence peripheral level (CP) (21a), and at the middle sentential level (21b):

- (21) a. [_{ForceP} ... [_{TopP} ... [_{FocP} ... [_{TopP} ... [_{IP} ... (Rizzi 1997)
 b. [_{CP} ... [_{IP} ... [_{TopP} ... [_{FocP} ... [_{TopP} ... [_{VP} [vP]]]]]]]]] (Belletti 2004)

I will also assume, in ForceP, a Speech Act Operator, which can have a morphological realization, as in Japanese *ka* for question, or an empty operator Σ_1 that results in a rising prosodic pattern, like in French, and Brazilian Portuguese. I also assume that in BP there may be two empty operators in ForceP, a Σ_1 for rising intonation, and a Σ_2 for falling intonation.

The most striking structural change that occurred in BP regarding *wh*-constructions was the appearance of *wh-in-situ* constructions as real questions, in the 20th century (Lopes-Rossi 1996). Until the 19th century, the FocP projection was exclusively in the sentential periphery, licensing both the V2 construction and the inverse-cleft *wh*-sentences.

In the 20th century, the higher FocP projection ceases to be projected, and a lower FocusP position left-adjacent to vP is activated (cf. Belletti 2004). The apparent ‘*wh-in-situ*’ in BP is proposed here, after Kato (2013a), to derive from a short movement of the *wh*-constituent to the vP adjacent FocP position as in (22a).¹³ It is important to stress that the construction that we are analyzing are real questions, with falling intonation (\Downarrow), determined by the operator Σ_2 in ForceP, and not echo-questions, which have a rising intonation (\Uparrow), determined by Σ_1 . The echo question in (22b), on the other hand, is analyzed as the result of the whole TP moved to Spec, ForceP (cf. Kayne 1994) interpreted as the sentence presupposition.

- (22) a. *As crianças dormem onde?* \Downarrow [genuine question]
 the children sleep where

13. Kato (2013a) was inspired by Miyagawa’s (2001) idea that Japanese has *wh*-movement of a short type, to a position in adjunction to T. My analysis proposes, however, that there is a designated position adjacent to vP in Belletti’s (2004) frame.

b. *As crianças dormem onde?* ↑ [echo question]

- (22') a. [_{ForceP} Σ₂ · [_{TP} as crianças dormem [_{FocusP} **onde** [_{VP} ~~as crianças~~ [_{VP} ~~dormem onde~~]]]]]]
- b. [_{ForceP} Σ₁ [_{TP} as crianças dormem **onde**] Σ₁ [_{TP} [_{TP} ~~as crianças dormem~~ [_{VP} ~~as crianças dormem~~ [_{VP} ~~dormem onde~~]]]]]]

Evidence that the fake *wh-in-situ* constructions have undergone movement is provided by Kato (2013a) using multiple adverbial adjuncts. In declarative clauses, multiple adverbial adjuncts come stacked at the end of the sentence. When multiple *wh*-adjuncts are moved, they appear in coordination. The fake *wh-in-situ wh*-words also appear in coordination, which indicates that they have been moved. Echo *in-situ* questions, on the other hand, can have only one *wh*-constituent *in-situ* (cf. (23e), with two independent intonations):

- (23) a. *Maria nasceu na Bahia em 1980.*
Maria was born in Bahia in 1980
- b. *Onde e quando Maria nasceu?*
where and when Maria was born
- c. *Maria nasceu onde e quando?* ↓
Maria was born where and when
- d. **Maria nasceu onde quando?* ↑
Maria was born where when
- e. *Maria nasceu onde?* ↑ *Quando?* ↑
Maria was born where when

But the fake *in-situ* questions like (24a) seem to compete with the outcome of *wh*-movement as in (24b). We should call the attention to the fact that there does not have to be a presupposition difference between the two forms in BP. This perfect functional equivalence can be seen in a verse from a Brazilian samba (24c):

- (24) a. *Você parou por que?*
you stopped why
- b. *Por que você parou?*
why you stopped
- c. *Porque (você) parou? (Você) Parou porque?*
why (you) stopped? (you) stopped why?
'Why did you stop?'

What remains to be answered regarding (24a) and (24b) is: do short and long *wh*-movement *grammatically* compete in BP vernacular? But recall that they are both innovations in the 20th century, and cannot be attributed to the instability of language change. We will discuss this point later, in Section 3.6.

3.5 From inverse-*wh*-clefts to canonic *wh*-clefts

At the same time that we have the appearance of the fake *wh-in-situ* questions, the inverse cleft-question like (25a), existent in both varieties since the end of the 17th century, starts to exhibit a competitive type of cleft in BP, the canonic cleft (25b), rarely found in 20th century corpora:

- (25) a. *Onde é que as crianças dormem?*
 where is that the children sleep [inverse cleft] (BP/EP 17th to 20th)
- b. *É onde que as crianças dormem?*
 is where that the children sleep
 ‘Where is it that the children sleep?’
 [canonic cleft *wh*-question] (BP 20th)

I analyze the new type, the canonic cleft, as a consequence of the activation of the lower FocP position in the copula clause, to which the *wh*-element moves:

- (25') b. [_{ForceP} Σ₂ [_{TP} É [_{FocP} **onde** [_{VP} é- [_{FiniteP} que [_{TP} as crianças dormem [_{VP} as crianças dormem [_{VP} ~~dormem onde~~]]]]]]]]]]

3.6 The 20th century innovative reduced cleft *Wh*-que

Previous analyses of the *reduced cleft* proposed that they exhibited a sort of “double filled Comp” as in (26)’ (cf. Mioto and Figueiredo Silva 1995; Hornstein, Nunes and Grohman 2005):

- (26) *Onde que as crianças dormiam?*
 where that the children slept
 ‘Where did the children sleep?’
- (26’) [_{CP} Onde [_C que [_{TP} as crianças dormiam [_{VP} ~~as crianças dormiam~~ [_{VP} ~~dormiam onde~~]]]]]]]

Kato and Raposo (1996) proposed instead that the reduced cleft was the reduction of the inverse cleft *wh*-question:

- (26)'' [_{FocP} Onde [_{Foc} (é) [_{TP} é- [_{VP} é- [_{CP} que [_{TP} as crianças dormiam [_{VP} ~~as crianças dormiam~~ [_{VP} ~~dormiam onde~~]]]]]]]]]

Kato (2007) claims, however, that the copula in BP can only be erased in sentence initial position:

- (27) a. *O seu cabelo *(é) lindo.*
 the your hair is beautiful

- b. (*É*) *lindo o seu cabelo.*
 is beautiful the your hair
 ‘Your hair is beautiful.’

A natural conclusion is that the reduced cleft does not derive from an inverse cleft, but from a canonic cleft, where the copula appears in sentence initial position.¹⁴ We can also claim that erasure was possible after the copula lost the *consecutio temporum*, becoming invariable, like with the inverse cleft construction:

- (28) a. *Era onde que as crianças dormiam?*
 was where that the children slept
 [Canonic cleft with *consecutio temporum*]
 b. *É onde que as crianças dormiam?*
 is where that the children slept
 [Grammaticalization of Tense]
 c. *É Onde que as crianças dormiam?*
 where that the children slept
 [Copula erasure at PF]
 ‘Where did the children sleep?’

The derivation of (28c) would be as follows:

- (28') c. [_{ForceP} ↓ Σ₂ [_{TP} (*É*) [_{FocusP} *onde* [_{VP} *é*- [_{FiniteP} *que* [_{TP} *as crianças dormiam*]]]]]]
 [_{VP} *as crianças dormiam* [_{VP} *dormiam onde*]]]]]

3.7 The appearance of *WH-SV*, the most recent pattern

The analysis presented in Kato and Raposo (1996) and Duarte and Kato (2002) linked the pattern *WH-SV* form to the loss of the Null Subject. While in Null subject languages the subject, when expressed, was postverbal in *wh*-questions, in BP the weakening of the inflectional system led to the creation of a paradigm of weak pronouns (cf. Kato 1999, Kato 2000) which could occupy Spec,T.¹⁵

14. Noonan (1989) considers a similar hypothesis, though in French the canonic cleft does not have the copula in initial position, because an expletive precedes it, but she mentions that (ii) is a disguised cleft.

(i) *C'est où que tá mis les oranges?*

(ii) *Où que tá mis les oranges.*

15. Ordoñez and Olarrea (2006) provide a similar explanation for Caribbean Spanish, which also lost the V2 pattern.

- (29) a. *Onde está ele dormindo?* EP
 where is he sleeping
- b. *Onde ele está dormindo?* BP
 where he is sleeping
 ‘Where is he sleeping?’
- (29′) a. [_{FocP} onde [está [_{TP} está [_{VP} ele está [_{VP} ~~está dormindo onde~~]]]]]
 b. [_{FocP} onde [[_{TP} ele está [_{VP} ele está [_{VP} ~~está dormindo onde~~]]]]]

However, the diachronic facts, and also the synchronic distribution of patterns, allow us to entertain a different hypothesis.

Recall that the V2 pattern was the oldest type in Portuguese history. While the pattern *Wh-SV* was the newest, appearing concomitantly with the pattern *Wh-que-SV*.

What is interesting, moreover, is that Brazilians avoid the Reduced Cleft (*WH-que-SV*) in written language while they prefer it in colloquial speech. In the written corpus,¹⁶ Kato and Mito (2005) found only one case of reduced cleft question, while in the oral corpus they found 18 cases.¹⁷ What Brazilians use instead, in written language, is the *WHSV* type, recognized by grammarians as the prototypical Brazilian norm.

The use of *WHSV* as a stylistic variant of *WH que SV* leads me to consider the possibility that from the reduced cleft type, a stylistic rule can erase the complementizer *que* in (30a–c), resulting in (30a′–c′) at PF. An interesting observation in Duarte (1992) was that the order *SV* appears first with adjunct *wh*-questions, and one can ask why. My claim is that with this type of questions, we have the ideal place to apply the stylistic rule of *haplology* (erase similar syllables).

- (30) a. *De que que os meninos riam?* (30a′)
 of what that the boys laugh
- a′. *De que os meninos riam?*
 of what the boys laugh
 ‘What were the boys laughing at?’
- b. *Para que que você quer isso?* (30b′)
 for what that you want this
- b′. *Para que você quer isso?*
 for what you want this
 ‘What do you want this for?’

16. The empirical, quantitative analysis was based on a *corpus* of newspapers found in the website <http://acdc.linguateca.pt/aceso/> for EP, the *subcorpus* Natura-Público and for BP the NILC-São Carlos.

17. For the spoken corpus, the authors used the spoken *corpus* *NURC*, and dialogues in comedies.

- c. *Por que que eles pararam?* (30c')
 why that they stopped
- c'. *Por que eles pararam?*
 why they stopped
 'What did they stop for?'

In cases where the *wh*-word is not similar to the complementizer *que*, we may suggest that erasure is the result of analogy with the adjunct cases.

Even the right dislocation structure in (31), with the order *wh*-VS, can result from the same grammatical structure, namely a canonic cleft sentence:

- (31) a. *Onde foi a Suzana?*
 b. (*É onde (que) (ela)foi a Suzana?*) [Right dislocation] BP
 'Where did Suzana go?'

4. Conclusions

The main structural change from OP to Modern BP was the change in the locus of FocusP projection:

– High Focus

- (32) a. *o que comprou a Maria?*
 what bought the Mary
 'What has Mary bought?'
 b. *O que é que a Maria comprou?*
 what is that the Mary bought
 'What is it that Mary bought?'
- (32') a. [_{ForceP} Σ₂ [_{FocP} ~~o que comprou~~ [_{TP} comprou [_{VP} a M. [_{VP} ~~comprou o que~~]]]] EP
 b. [_{ForceP} Σ₂ [_{FocP} o que é [_{TP} é- [_{VP} é- [_{CP} **que** [a M. comprou ~~o que~~ ?]]]]] EP/BP

– Medial/low Focus

- (33) a. *A Maria comprou o que?*
 the Mary bought what
 'What did Mary buy?'
 b. *É o que que a Maria comprou?*
 Is what that the Mary bought
 'What was it that Mary bought?'

- (33') a. $[_{\text{ForceP}} \Sigma_2 [_{\text{TP}} \text{a Maria comprou } [_{\text{FocP}} \text{o que } [_{\text{VP}} \text{a M. } [_{\text{VP}} \text{comprou o que}]]]]]]$ BP %EP
- b. $[_{\text{ForceP}} \Sigma_2 [_{\text{TP}} (\acute{\text{e}}) [_{\text{FocP}} \text{o que } [_{\text{VP}} \acute{\text{e}} [_{\text{CP}} \text{que } [_{\text{IP}} \text{a M. comprou o que}]]]]]]]]$ BP *EP

Figure 3 summarizes the conclusions of the present chapter.

Long <i>wh</i> -movement		Short <i>wh</i> -movement			
V2	Inverse cleft	Fake <i>wh</i> -in-situ	Canonic cleft		
<i>wh</i> -VS	<i>wh</i> -é que-SV	SV <i>wh</i> -	É <i>wh</i> -que-SV	<i>wh</i> -que-SV	<i>wh</i> -SV
OP CIP EP	CIP EP BP	BP			

Figure 3. From Old and Classic Portuguese to BP

The only case with long *wh*-movement in BP is the inverse cleft, which we consider a residue of the old grammar, while it is the most vernacular type in EP.

I finally conclude that in BP vernacular there is no long *wh*-movement, nor a real *wh*-in-situ construction except in echo questions. We only have a short *wh*-movement in all cases. All the other variants result from phonological reduction and, therefore, do not produce syntactic “doublets” but stylistic variants. We follow Chomsky and Lasnik (1977) here, for whom stylistic rules occur at PF, with no effect on meaning.

Acknowledgements

This work had the support of CNPq grant 305515/2011–2017. I thank the two reviewers for the important and generous contributions on form and content of this article. I also thank Marcello Marcelino for his contribution on the first version of the paper.

References

- Ambar, Maria Manuela. 1988. “Para uma sintaxe da Inversão Sujeito verbo em Português.” Ph.D. dissertation, University of Lisbon.
- Ambar, Maria Manuela. 2003. “*Wh*-Asymmetries.” In *UG and External Systems: Language, brain and computation*, ed. by Ana Maria Di Sciullo, 209–249. Amsterdam: John Benjamins.
- Aronoff, Mark. 1976. *Word Formation in Generative Grammar*. Cambridge: MIT Press.
- Barbosa, Pilar. 2001. “On Inversion in *Wh*-questions.” In *Romance Inversion*, ed. by Aafke Hulk, and Jacques-Yves Pollock, 2–59. New York: Oxford University Press.
- Belletti, Adriana. 2004. Aspects of the low IP area.” In *The Structure of IP and CP: The Cartography of Syntactic Structures*, ed. by Luigi Rizzi, 16–51. New York: Oxford University Press.

- Britto, Helena. 2000. "Syntactic codification of categorical andthetic judgments in Brazilian Portuguese." In *Brazilian Portuguese and the Null Subject Parameter*, ed. by Mary A. Kato, and Esmeralda V. Negrão. 195–222. Frankfurt: Vervuert/Iberoamericana.
- Cheng, Lisa Lai-shen, and Johan Rooryck. 2000. "Licensing *wh*-in-situ." *Syntax* 3: 1–19. <https://doi.org/10.1111/1467-9612.00022>
- Chomsky, Noam, and Howard Lasnik. 1977. "Filters and Control." *Linguistic Inquiry* 8: 425–504.
- Duarte, Maria Eugenia Lamoglia. 1992. "A perda da ordem V(erbo) S(ujeito) no português do Brasil." *DELTA* 8: 37–52.
- Duarte, Maria Eugenia, and Mary A. Kato. 2002. "A Diachronic Analysis of Brazilian Portuguese *Wh*-Questions." *Santa Barbara Portuguese Studies* VI: 326–339.
- Hornstein, Norbert, Jairo Nunes, and Kleantes Grohman. 2005. *Understanding Minimalism*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511840678>
- Huang, C.-T. James. 1982. "Logical Relations in Chinese and the Theory of Grammar." Ph.D. dissertation, MIT.
- Kato, Mary A. 1987. "Inversão da ordem SV em interrogativas no português." *DELTA* 3 (2): 243–252.
- Kato, Mary A. 1993. "The distribution of null and pronominal objects in Brazilian Portuguese." In *Linguistic Perspectives on the Romance Languages: Selected Papers from the XXI Linguistic Symposium on Romance Languages*, ed. by William J. Ashby, Marianne Mithun, and Giorgio Perissinotto, 225–235. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.103.23kat>
- Kato, Mary A. 1999. "Strong pronouns and weak pronominals in the null subject parameter." *Probus* 11 (1): 1–37. <https://doi.org/10.1515/prbs.1999.11.1.1>
- Kato, Mary A. 2000. "The partial pro-drop nature and the restricted VS order in Brazilian Portuguese." In *Brazilian Portuguese and the Null Subject Parameter*, ed. by Mary A. Kato, and Esmeralda V. Negrão. 223–258. Frankfurt: Vervuert/Iberoamericana.
- Kato, Mary A. 2007. "Free and dependent small clauses in Brazilian Portuguese." *DELTA* 23: 85–111. <https://doi.org/10.1590/S0102-44502007000300007>
- Kato, Mary A. 2013a. "Deriving *wh*-in-situ through movement." In *Information Structure and Agreement*, ed. by Victoria Camacho-Taboada, Ángel L. Jiménez-Fernández, Javier Martín-González, and Mariano Reyes-Tejedor, 175–191. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.197.06kat>
- Kato, Mary A. 2013b. "*WH*-movement in the history of Brazilian Portuguese." *STIL* 5: 59–76. http://www.ciscl.unisi.it/doc/doc_pub/
- Kato, Mary A. 2014. "Focus and *WH*-questions in Brazilian Portuguese." In *Linguistic variation: Confronting fact and theory*, ed. by Nathalie Dion, André Lapierre, and Rena Torres Cacoullous, 111–130. New York: Routledge.
- Kato, Mary A., and Carlos Mioto. 2005. "A multi-evidence study of European and Brazilian *wh*-questions." In *Linguistic evidence: empirical, theoretical and computational perspectives*, ed. by Stephan Kepser, and Marga Reis, 307–328. Berlin: Mouton De Gruyter. <https://doi.org/10.1515/9783110197549.307>
- Kato, Mary A., and Eduardo Raposo. 1996. "European and Brazilian word order: questions, focus and topic constructions." In *Aspects of Romance Linguistics*, ed. by Claudia Parodi, Carlos Quicoli, Mario Saltarelli, and Maria Luisa Zubizarreta, 267–277. Washington, D.C.: Georgetown University Press.
- Kato, Mary A., and Ilza Ribeiro. 2009. "Cleft sentences from Old Portuguese to Modern Brazilian Portuguese." In *Focus and Background in Romance Languages*, ed. by Andreas Dufter, and Daniel Jacob, 123–154. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.112.06kat>

- Kato, Mary A., and Fernando Tarallo. 2003. "The loss of VS syntax in Brazilian Portuguese." In *Dialogue between Schools: Sociolinguistics, conversational analysis and generative theory in Brazil*, ed. by Brigitt Schliebe-Lange, Ingedore Koch, and Konstanze Jungbluth, 101–129. Münster: Nodus Publicationen, Klaus D. Ditz.
- Kayne, Richard. 1994. *The Antysymmetry of Syntax*. Cambridge, MA: MIT Press.
- Kayne, Richard. 2001. Pronouns and their antecedents. In S. Epstein and D. Seely (orgs) *Derivation and explanations in the Minimalist program*. Malden, MA: Blackwell. 133–166.
- Kayne, Richard, and Jacques-Yves Pollock. 1978. "Stylistic Inversion, Successive Cyclicity and Move NP in French." *Linguistic Inquiry* 9: 595–621.
- Kroch, Anthony. 1994. "Morphosyntactic variation." In *Papers from the 30th Regional Meeting of the Chicago Linguistics Society: Parasession on Variation and Linguistic Theory*, ed. by Katherine Beals, 180–201. Chicago: CLS.
- Kuroda, S-Yuki. 1972. "The categorial and the thetic judgment: Evidence from Japanese Syntax." *Foundations of Language* 9: 153–185.
- Lopes-Rossi, Maria Aparecida. 1996. "A Sintaxe Diacrônica das Interrogativas-Q do Português." Ph.D. dissertation, UNICAMP.
- Martins, Ana Maria. 1994. "Clíticos na História do Português." Ph.D. dissertation, University of Lisbon.
- Mioto, Carlos, and M. Cristina Figueiredo Silva. 1995. "Wh que = Wh é que?" *DELTA* 11 (2): 301–311.
- Miyagawa, Shigeru. 2001. "EPP, Scrambling, and *wh-in situ*." In *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 293–338. Cambridge, MA: MIT Press.
- Noonan, Maire. 1989. "Operator licensing and the case of French interrogatives." In *Proceedings of the 8th West Coast Conference on Formal Linguistics*, ed. by Jane Fee, and Kathryn Hunt, 315–330. Stanford: CSLI.
- Ordoñez, Francisco, and Anton Olarrea. 2006. "Microvariation in Caribbean/non-Caribbean Spanish." *Probus* 18 (3): 59–97.
- Rizzi, Luigi. 1997. "The fine structure of the left periphery." In *Elements of Grammar*, ed. by Liliane Haegeman, 281–337. Dordrecht: Kluwer.
https://doi.org/10.1007/978-94-011-5420-8_7
- Uriagereka, Juan. 1995. "An F position in Western Romance." In *Discourse Configurational Languages*, ed. by Katalin É. Kiss, 153–175. New York: Oxford University Press.
- Vergnaud, Jean-Roger, and Maria Luisa Zubizarreta. 2005. "The representation of focus and its implications: Towards an alternative account of some 'intervention effects.'" In *Organizing Grammar- Linguistic Studies in Honor of Henk van Riemsdijk*, ed. by Hans Broekhuis, Norbert Corver, Reny Huybregts, Ursula Kleinhenz, and Jan Koster, 641–60. Berlin: Mouton de Gruyter.

Clitic doubling, person and agreement in French hyper-complex inversion

Richard S. Kayne
New York University

The paper claims that Hyper-Complex Inversion (HCI) is an instance of clitic doubling that is subject to a person restriction not found with more familiar cases of clitic doubling. This restriction is argued to result from an incompatibility between the postverbal subject clitic (SCL) of HCI and the demonstrative structure associated with first and second person pronouns. That demonstrative structure also plays a role in asymmetries that hold concerning the possible silence, in certain cases, of third person pronouns, but not first or second person pronouns. HCI shares with past participle agreement the property that it is incompatible with an unmoved lexical direct object, in a way that presents a challenge to Agree, if Agree is taken to be available even in the absence of movement. The SCL of HCI sentences can sometimes climb out of an embedded infinitive, in a way related to the Comp-area character of its landing site.

Keywords: doubling, agreement, inversion, person, demonstrative

1. HCI

Standard French has, in root interrogatives:¹

- (1) Est-il heureux? ('is he happy')

This inversion applies only if the subject is a pronominal clitic:²

- (2) *Est Jean heureux? ('is J happy')

1. In what follows, 'Standard French' is abbreviated as 'French'.

2. French has another, distinct inversion construction informally called 'stylistic inversion' that sometimes (but not in yes-no questions) overlaps with subject clitic inversion. See Kayne and Pollock (2001) and references cited there.

French also allows a variant of (1) in which a non-dislocated preverbal subject co-occurs with the post-verbal pronominal subject clitic. Kayne (1972) informally called this ‘Complex Inversion’ (henceforth CI):

- (3) *Cela est-il vrai?* (‘that is it true’ = ‘is that true?’)

In the appropriate register, CI, as in (3), is productive. Relevant to this paper is the fact that CI is compatible with object clitics (henceforth OCLs):

- (4) *Cela la gêne-t-il?* (‘that her bothers it’ = ‘does that bother her?’)

Central to this paper is an important observation due to Morin (1985: 796), namely that alongside (sometimes instead of)³ (4) there are speakers who accept, with the same interpretation as (4):⁴

- (5) *Cela la gêne-t-elle?* (‘that her bothers she’ = ‘does that bother her?’)

In (5) the postverbal nominative subject clitic (here *elle*) agrees in gender (and number) with the preverbal accusative OCL (here *la*). This contrasts with ordinary CI, as in (3) and (4), in which the postverbal subject clitic (there *il*) agrees with the preverbal subject *cela*. I shall use for (5) the informal term Hyper-Complex Inversion (henceforth HCI).

2. HCI as clitic doubling

In many ways HCI and CI are similar,⁵ so that one can think of HCI as a subclass of CI characterized by the agreement, in HCI, between subject clitic (henceforth

3. See note 21. There are speakers (e.g. Anne Zribi-Hertz, p.c.) who strongly reject HCI; cf. the phenomena concerning *tous* (‘all’) that are discussed in Kayne (1975, sect. 1.11), which meet with (sometimes strong) disagreement across speakers. In this paper, I will not attempt to delineate the parameter(s) underlying acceptance vs. non-acceptance of HCI (there may be a link to past participle agreement, which is also not uniformly accepted in French; cf. also sect. 7 below).

4. For further background, see Kayne and Pollock (2012, 2014), from which early sections of this paper draw freely. HCI sentences are best when the lexical subject contains no lexical noun, as with *cela* (‘that there’); the HCI examples in the text have feminine SCL *elle*, but there are also acceptable examples of HCI that have masculine SCL *il* agreeing with masculine OCL *le* in the presence of a feminine lexical subject:

- (i) *Cette affaire/chose le gêne-t-il?* (‘that business/thing him bothers t it’ = ‘does that business/thing bother him?’)

5. For example, both are restricted to root contexts lacking any complementizer, both are limited to interrogatives and some affective contexts, both have the property that the postverbal subject pronoun must be a clitic. For additional details on CI, see Pollock (2006).

SCL) and OCL (and by the non -agreement, in HCI, between SCL and preverbal subject).⁶ Both CI and HCI can be thought of as instances of clitic doubling, i.e. as having something in common with the well-known dative clitic doubling found across Spanish. One key difference is that CI and HCI centrally involve not dative clitics but rather nominative ones (SCLs).⁷

Let me now adopt the ‘complex constituent’ approach to French CI clitic doubling proposed in Kayne (1972, sect. 3).⁸ Thus in a CI example like (4) the DP *cela* and the SCL *il* start out within a phrase (a complex DP, in more recent terms) that excludes the verb (and the object):⁹

(6) ...[cela il] gêne la

By extension, in HCI examples like (5), the SCL *elle* starts out paired with the OCL, rather than with the lexical subject:

(7) ...cela gêne [la elle]

The agreement seen in CI/HCI is in this way reduced to agreement (in gender and number) within a complex DP. (In both (4) and (5) the complex DP is split apart in the course of the derivation.)

Contrary to CI, simple SCL inversion (henceforth SCLI) of the sort seen earlier in (1) has no HCI-like counterpart that would correspond to (5). Note first that SCLI is compatible with an OCL:

(8) La gêne-t-il? (‘her bothers it/he’ = ‘does he/it bother her?’)

Here *la* and *il* correspond to distinct arguments. If there existed an HCI-like counterpart, then agreement between SCL and OCL would be possible, keeping the interpretation constant. That is not, however, possible in (8). The following is well-formed, but not with the interpretation of (8):

6. There also exist instances of HCI in which the SCL agrees with a preposed wh-phrase; these will be set aside in this paper, as will the marginal cases of HCI in which the SCL agrees with a dative clitic (cf. Morin 1985: 796); on the latter, see also Kayne and Pollock (2014: 100).

7. On the nominative status of French subject clitics, see Napoli (1981) and Kayne and Pollock (2001, sect. 5).

8. Cf. Uriagereka (1995: 81) on Spanish.

9. This complex DP approach to clitic doubling has something in common with Szabolcsi’s (1983, 1994) proposal for simple possession sentences like (the Hungarian counterpart of) *John has a sister*, according to which (by transposition to English) the possessor *John* originates within a DP containing *a sister*. Cf. Kayne (1993) and Boneh and Sichel (2010).

- (9) La gêne-t-elle? ('her bothers she/it' = 'does she/it bother her?')¹⁰

The reason that (9) cannot be related to (8) in the way that (5) is related to (4) is the following. In (5), *elle* can be taken to be a double of *la* (both then being part of the object argument), since there is still *cela* to fill the role of subject argument. Whereas in (9), if we were to take *elle* to be a double of *la*, there would be nothing left to fill the role of subject argument.¹¹

3. A restriction concerning SCLs

When the preverbal subject is itself a SCL, neither CI nor HCI is possible.¹² Thus alongside (10) with SCL *ils*, French allows SCLI, as in (11):

- (10) Ils la voient. ('they her see')

- (11) La voient-ils?

But it does not allow either CI (12) or HCI (13):

- (12) *Ils la voient-ils?

- (13) *Ils la voient-elle?

4. The *-t-* morpheme

The sharp deviance of (12) and (13) can be thought of as reflecting the fact that French cannot license two SCLs in one simple sentence, as opposed to French being able, in CI/HCI, to license one (postverbal) SCL and one (preverbal) lexical subject DP at the same time.

Setting aside left- and right-dislocation, we can see that such double licensing is possible only if the SCL is postverbal. This is shown using CI in the following pair of examples:

10. A counterpart of this example appears to be possible in the North Italian dialect (close to Paduan) discussed by Penello (2003, 2007 (11b)). The contrast with French might be related to that dialect's being a partial pro-drop language; see also Roberts (1993), Pollock (2006) and Roberts (2010: 119).

11. Since French is not a null subject language of the Italian sort. The text proposal is in the spirit of Morin (1985: 796).

12. As opposed to the dialect studied by Penello (2003, 2007). The *-ti* of popular French mentioned by Morin (1985, 794) is also compatible with preverbal subject clitics.

(14) *Cela est-il vrai?* ('that is it true')

(15) **Cela il est vrai.*

Similarly, for HCI we have:

(16) *Cela la gêne-t-elle?* ('that her bothers t she' = 'does that bother her?')

(17) **Cela elle la gêne.*

Let us take (15) and (17) to be excluded for the following reason. Preverbal SCLs and preverbal non-dislocated lexical subject DPs are, in French,¹³ Case-licensed by a functional head that can license only one of them in a given simple sentence.

If so, then postverbal SCLs, as in (14) and (16), must have access to an extra licenser, one that is not available to preverbal subjects of any type. In the spirit of Pollock (2006), let us take the *-t-* morpheme of CI and HCI (seen clearly in (16)) to play a key role in licensing the postverbal subject clitic.¹⁴ This is plausible since there is without exception a pronounced *-t-* immediately preceding the postverbal SCL in both CI and HCI.¹⁵

5. Remnant movement and *-t-*

Thinking of the limitation of this *-t-* to root contexts, of a partial similarity to Germanic complementizer agreement,¹⁶ and of Shlonsky (1994), let us take *-t-* to be a morpheme located above IP, somewhere in the Comp area, in Rizzi's (1997) sense. Of importance now are two properties of *-t-*. The first, already briefly discussed, is that it participates in the Case-licensing of the following SCL (found in the projection just below it).¹⁷

13. As opposed to various dialects in northern France and in northern Italy; on the latter, see Poletto (2000) and references cited there. On (preverbal) SCLs, see also Kayne (1983).

14. Cf. also Schoorlemmer (2006).

15. A more syntax-friendly French orthography would arguably write (14) as:

(i) *Cela est-t-il vrai?*

in which only one *t* would be pronounced, in a way consistent with general properties of French phonology.

16. For discussion, see, for example, Gruber (2008). Why Germanic 'complementizer agreement' is (apparently) limited to the 'OV' Germanic languages needs to be accounted for – see Kayne (1994: 52) and Koopman (2005a, note 25).

17. The licensing of the post-*t* SCL may in addition depend on finite verb agreement; for relevant discussion, see Kayne and Pollock (2014, sect. 7 and 8), who broach the possibility of

The second important property of this *-t-* is that it attracts to its Spec a phrase containing the lexical subject plus the finite verb (along with any intervening OCLs). In (16), for example, *-t-* attracts the phrase ‘[*cela la gène*]’, as indicated in the following derivation:

- (18) *cela gène* [la elle] → OCL movement (pied-piping the SCL)¹⁸
cela [la elle]_i *gène* t_i → raising of SCL¹⁹
 elle_j *cela* [la t_j]_i *gène* t_j → merger of *-t-*
 t elle_j *cela* [la t_j]_i *gène* t_j → remnant IP movement
 [cela [la t_j]_i *gène* t_j]_k t elle_j t_k

with the last step involving remnant IP movement essentially as in Pollock (2006).²⁰

6. HCI and clitic climbing

All the CI and HCI examples given so far that contain an OCL have had that OCL (*la*) preceding the SCL (*il* or *elle*):²¹

- (19) *Cela la gène-t-il?* = CI
 (20) *Cela la gène-t-elle?* = HCI

CI readily allows a SCL to precede an unrelated OCL, given some embedding:

there being two agreements, in at least some cases. Licensing (in part) by finite verb agreement would reinforce the nominative character of these SCLs (see note 7); as a reviewer notes, if they were not strictly nominative, one would wonder why there is no counterpart to CI/HCI with a morphologically accusative post-*t* clitic.

18. On this step, see in part Kayne (2002, sect. 9). The subsequent SCL-raising step recalls Caha (2010).

19. The raising of SCL across *cela* leads to a relativized minimality question. It may be that SCL and DP count as sufficiently different (which might lead to an alternative account of the double SCL restriction discussed earlier, if moving one SCL across another is prohibited).

20. For HCI sentences in which the SCL is followed by other material, as in:

- (i) *Cela la remplit-elle de joie.* (‘that her fills-she of joy’ = ‘does that fill her with joy?’)

that material, here *de joie*, will need to be scrambled out prior to the remnant movement in question, much as in many derivations in Koopman and Szabolcsi (2000), though there’s some tension with the use to which such scrambling was put in Kayne (1998).

21. Left open in this paper is the question why some speakers accept only the HCI variant of such pairs (cf. Morin 1985, note 13). This may be related to the fact that Dominique Sportiche, who accepts (p.c.) both (19) and (20), finds the HCI variant less elevated than the CI one.

- (21) *Cela va-t-il la gêner?* ('that is-going-to it her bother' = 'is that going to bother her?')

In (21), SCL *il* precedes OCL *la*, with which it does not agree and with which it is derivationally unrelated.

The question arises as to whether in a configuration like that in (21), HCI would be possible, i.e. whether or not a SCL can ever agree with an OCL that follows it. Morin (1985, 796) says no, but some speakers find acceptable some sentences such as:²²

- (22) *Cela va-t-elle la déranger?* ('that is-going-to *t* she her disturb' = 'is that going to disturb her?')

whose CI counterpart, without agreement between *il* and *la*, is:

- (23) *Cela va-t-il la déranger?*

HCI examples such as (22) are not, however, possible if the OCL is within a finite embedding (with the SCL in the matrix, as usual):

- (24) *Cela implique-t-il que Jean la voit souvent?* ('that implies it that J her sees often')
- (25) **Cela implique-t-elle que J la voit souvent?*

22. Especially in a CLLD (clitic left dislocation, as in Cinque (1990)) context:

- (i) *Cette personne, cela va-t-elle la déranger?* ('this person, ...')

which suggests the involvement of a pied-piping-like movement of *cette personne*.

Having the SCL agreeing with a following OCL is sometimes felt to be less good in the plural:

- (ii) ^{??}*Cela va-t-elles les déranger?* ('that is-going-to *t* they them disturb' = 'is that going to disturb them')

This may be related to the fact that for some speakers a plural OCL in HCI calls for plural verb agreement –cf. Kayne and Pollock (2014, sect. 7). Possibly, there is also link to the fact that Spanish *leísmo* is less widespread in the plural than in the singular, as noted in Navarro and Neuhaus (2016: 80, 83). A reviewer also suggests a possible link to the fact that in (much) Catalan past participle agreement in the plural seems to be dependent on overt agreement in gender; cf. Bonet (1991, 165n).

In raising past the OCL *la* that it agrees with, the SCL *elle* in (22) has something in common with the Italian agreeing past participle *offerta* in the following example (from Longobardi 1985, note 23):

- (iii) *Offerte a sua moglie, credo che Mario ancora non le abbia.* ('offered to his wife, I-believe that M still not them has' = 'I believe that M still hasn't offered them to his wife')

in which *offerta* has, via remnant movement, moved past the OCL *le* that it agrees with.

(24) is a well-formed CI example, in which SCL *il* agrees with subject *cela*. (25) shows that trying to turn (24) into an HCI example by having SCL *elle* agree with the following OCL *la* is not possible, contrary to (22) (for the relevant speakers).

More strikingly, HCI is possible to one degree or another with an infinitival embedding in the manner of (22) only with matrix verbs/predicates of the ‘restructuring’ type. Thus, alongside (22) and the similar:

- (26) ?*Cela pourrait-elle la gêner?* (‘that could she her bother’ = ‘could that bother her?’)

in which HCI is to some extent available, we have the fact that the following well-formed CI example:

- (27) *Cela a-t-il l’air de la gêner?* (‘that has t it the air of her to- bother’ = ‘does that look like it bothers her?’)

has no well-formed HCI counterpart:

- (28) **Cela a-t-elle l’air de la gêner?*

The similarity holding here between HCI and, say, Italian object clitic climbing (with respect to sensitivity to ‘restructuring’) suggests that the SCL *elle* in (22) and (26) must have raised into the matrix from within the infinitive in French in a way parallel to OCL clitic climbing in Italian.

The derivation of (22), for example, will (for those speakers who accept it) be approximately as in (29) (cf. the derivation given in (18)):

- (29) déranger [la elle] → OCL movement (pied-piping the SCL)
 [la elle]_i déranger t_i → merger of matrix *va* and of subject *cela*
cela va [la elle]_i déranger t_i → scrambling of infinitive phrase²³
 [[la elle]_i déranger t_i]_j *cela va* t_j → raising of SCL and merger of -*t*-
*t elle*_k [[la t_k]_i déranger t_i]_j *cela va* t_j → remnant IP movement
 [*cela va* t_j] t elle_k [[la t_k]_i déranger t_i]_j

This yields (22), repeated here:

- (30) *Cela va-t-elle la déranger?* (‘that is-going-to *t* she her disturb’ = ‘is that going to disturb her?’)

The SCL raising seen in (29) must be available only when the matrix predicate is of the restructuring type, not otherwise. A more general formulation is:

23. Cf. Collins (2005) on ‘smuggling’, which interacts here with the relativized minimality question mentioned in note 19; for a partial precursor of smuggling, see Kayne (1975: 272, 329); also, Kayne (1994: 54) on nominative anaphors.

- (31) Only in the case of restructuring predicates can pronominal clitics raise out of infinitival complements.

This statement is intended to hold even if, as in (29), the infinitive phrase has previously scrambled.²⁴

The formulation in (31) leaves open the curious fact that in (29)/(30) the SCL *elle* has succeeded in escaping from the infinitive phrase despite French not normally (apart from causatives) allowing OCLs to escape from infinitive phrases, even those embedded under restructuring predicates,²⁵ as seen in the contrast between (30) and (32):

- (32) *Cela la va-t-elle déranger?

Continuing to think in terms of the derivation (29), the key difference between SCL *elle* in (30) (and (32)) and OCL *la* in (32) may lie in the fact that the landing site of SCL-raising in these HCI inversion derivations is up in the Comp area in Rizzi's (1997) sense, i.e. above the normal (preverbal) position of the subject, as seen in both (18) and (29), whereas the landing site of OCLs is in French invariably below normal subject position.²⁶

Another way of putting this is to say that SCL-raising in HCI derivations is A-bar-like, whereas OCL movement is not:

- (33) Raising out of infinitival phrases (of the sort that crosses a subject position)²⁷ is possible in French with A-bar-like movements only.

24. If SCL-raising were to precede infinitive phrase scrambling, then by the extension condition the infinitive phrase would, incorrectly, end up preceding the SCL in (30). Alternatively, it might be possible to rework (29) in the manner of Chomsky's (2008) discussion of CED effects.

25. For relevant discussion, see Kayne (1989a, 1991).

26. One will need to bring in Portuguese OCLs here; for relevant discussion, see Uriagereka (1995).

27. This is to allow for subject-to-subject raising and for raising of an ECM subject; see Pollock (1978, 1985). It will also allow for OCL-raising out of infinitives in (certain) causatives; cf. Kayne (1975, Chapters 4 and 6) and Rouveret and Vergnaud (1980). Also, see Kayne (1981) on the extra possibilities for the movement of *tout* (vs. OCLs); the fact that moved *tout* ('all') doesn't license HCI, as seen in:

- (i) *Cette affaire gêne-t-il tout? ('that affair upsets *t* it everything')

might be related to moved *tout* not licensing complementizer-like *qui*, as discussed in that paper. On the fact that *tout* has moved in (i), see Pollock (1989, note 7) and, for Italian, Cinque (1995, Chapter 9).

7. A familiar problem for Agree

Of further note is the contrast between (30) and the following:

- (34) **Cela va-t-elle déranger Marie?* ('that is-going-to *t* she disturb Mary' = 'is that going to disturb Mary?')

In both (30) and (34) the agreeing SCL *elle* precedes what it agrees with (*la, Marie*). Yet only in (30) is the result acceptable, indicating that derivation-final word/morpheme order is not what is at issue.

Rather the contrast between (30) and (34) should be reduced to that holding between the following two simpler cases:²⁸

- (35) *Cela la dérange-t-elle?* ('that her bothers *t* she' = 'does that bother her')
 (36) **Cela dérange-t-elle Marie?* (= 'does that bother Mary?')

The agreeing SCL that characterizes HCI can only successfully agree, as it (*elle*) does in (35), with a direct object if that direct object (*la* in (35)) has moved leftward (to a sufficiently high position, including in (30)). In (36), the object *Marie* has either not moved at all, or else has not moved high enough to license SCL-agreement of the HCI sort.

The contrast between (35) and (36) strongly recalls a basic property of French and Italian past participle agreement,²⁹ as illustrated in French by:

- (37) *Jean l'a repeinte.* ('J it(fem.) has repainted(fem.) = 'J has repainted it')
 (38) *Jean a repeint/*repeinte la table.* ('J has repainted (masc.)/(fem.) the table')

In (37), the direct object clitic *la* (which here loses its *-a*) has moved up past the auxiliary; the past participle *repeint* agrees in gender (and number) with that *la*. In (38), on the other hand, the direct object *la table* has not moved (far enough) up and agreement is impossible.

The contrast seen in (37) vs. (38) is unexpected if Agree need not be associated with movement (and if Agree is taken to underlie past participle agreement).³⁰ The same would hold for HCI if one took Agree to underlie the agreement found in HCI sentences. One might pursue that Agree possibility by having *-t-* in (35) act

28. As pointed out by Morin (1985: 796).

29. Cf. Kayne (1985, 1989b) and Belletti (2006). As Baker (2008: 198, note 30) notes, this upward bias for past participle agreement poses a problem for his characterization of agreement in Indo-European languages.

30. Cf. Kayne (2008, to appear) for an analysis of expletive *there* that does not need movement-less Agree; also, Koopman (2003, 2005b). On Agree, see Chomsky (2000, 2001).

as a probe for *la*, inducing remnant IP-movement as in (29), with *elle* then being the spellout of the agreement relation. But in that case the impossibility of the agreement shown in (36) would be unexpected, if Agree could be dissociated from movement, insofar as *-t-* in (36) could find *Marie* as goal.

8. Pronominal clitics vs. agreement morphemes

On the other hand, one could try to maintain the availability of movement-less Agree in the face of (36) (though (38) would remain a challenge) by denying that Agree is relevant to HCI at all. That would in all likelihood lead to denying more generally that Agree is relevant to clitic doubling (which would diminish the interest of Agree), and would in all likelihood lead to saying that there is a sharp difference between clitic doubling and agreement. Whether there is such a sharp difference is a question that can be asked independently of Agree. Let me now turn briefly to that question.

The kind of agreement seen in (35), in which SCL *elle* agrees with feminine singular OCL *la*, does differ sharply from more familiar instances of agreement in French, which otherwise disallow *elle* as the spellout of feminine singular agreement. Thus in (37) the past participle agreement morpheme is *-e* and cannot be *elle*:

(39) *Jean l'a repeintelle.

Similarly, DP-internal adjective or indefinite article agreement in French shows *-e* for feminine singular (*grand+e*, *un+e*):

(40) une grande maison ('a(fem.) big(fem.) house(fem.)')

and cannot show *elle* instead:

(41) *une grandelle maison; *unelle grande maison; *unelle grandelle maison

Conversely, feminine singular *-e* cannot replace *elle* in (35) or in any other example of HCI:

(42) *Cela la dérange-t-e?

The same holds for CI:

(43) Marie a-t-elle une grande maison? ('M has -t- she a big house')

(44) *Marie a-t-e une grande maison?

Following a long tradition, I take the postverbal SCL *elle* in question (in CI, in HCI, and also in SCLI (11)) to be a pronominal clitic, and the *-e* of (40) and

(37) not to be a pronominal clitic. In French, this distinction goes with a difference in form. Third person non-reflexive pronominal clitics always contain an *l*, as seen in SCLs in:

(45) *il* (m.sg.), *ils* (m.pl.), *elle* (f.sg.), *elles* (f.pl.)

in accusative OCLs:

(46) *le* (m.sg.), *la* (f.sg.), *les* (pl.)

and in dative OCLs:

(47) *lui* (sg.), *leur* (pl.)

whereas the *-e* of (40) and (37) does not contain an *l*.

9. Person and *l*

The pronominal clitic status of SCL *elle* in HCI examples like (35), repeated here:

(48) *Cela la dérange-t-elle?* ('that her bothers t she' = 'does that bother her')

combined with the pronominal clitic status of OCL *la* in the same example leads to the unsurprising conclusion that HCI (like CI) is to be thought of as an instance of clitic doubling.

Conversely, since *-e* is not a pronominal clitic, past participle agreement examples like (37), repeated here:

(49) *Jean l'a repeinte.* ('J it(fem.) has repainted(fem.)')

are not instances of clitic doubling.

It is important to note, however, that all cases of clitic doubling themselves involve agreement. In particular, and without exception as far as I know, the following holds:³¹

(50) Clitic doubling invariably shows person agreement between the clitic and the other element or phrase in question.

31. Colloquial Spanish allows number agreement not to hold with third person dative clitic doubling; see Butt and Benjamin (1988, sect. 11.14.3). As a reviewer points out, the dative clitic must, however, agree in CLLD (Cinque 1990) sentences, and also, though less sharply, in 'V PP DP' sentences.

As a second reviewer points out, Zagana (2002: 68) gives an example lacking gender agreement. Instances of non-agreement in person have not yet been discovered, that I know of.

This is true of HCI, as in (48) (in which the other element is a second clitic). It is true of CI. It is true of Spanish clitic doubling, both of the dative sort and of the (less widely found) accusative sort.³²

What this means is that proposals to distinguish clitic doubling from agreement, as, for example, in Preminger (2009), must be understood, given (50), as shorthand for distinguishing clitic doubling (which itself involves agreement) from instances of agreement that do not involve pronominal clitics.³³

The difference between clitic doubling, on the one hand, and non-clitic-doubling agreement, on the other,³⁴ manifests itself in a striking way in French. There is a clear difference between HCI (an instance of clitic doubling) and past participle agreement (not an instance of clitic doubling) that involves person, in a certain way. In French a past participle can agree in gender (and number) with a first or second person pronoun accusative OCL:

(51) Jean t'a prise par le bras. (French: 'J you(fem.) has taken(fem.) by the arm')

In contrast, while CI in French can readily have a first or second person OCL, as in:

(52) Cela te gêne-t-il? ('that you bothers it' = 'does that bother you?')

HCI cannot.³⁵ Even if the OCL in (52) is understood to be feminine, the SCL must remain *il* (pairing with *cela*); this *il* cannot be replaced by feminine *elle* (which would be agreeing in gender with *te*):

(53) *Cela te gêne-t-elle?

32. On Spanish *Nos vio a los lingüistas* ('us (s)he-saw to the linguists' = '(s)he saw us linguists'), which almost certainly contains a silent first person plural non-clitic pronoun, see Torrego (1996: 124) and Ordóñez and Treviño (1999); also, Kayne (2009).

33. Preminger's (2009) use of intervention effects as a tool for distinguishing clitic doubling from agreement will need to be recalibrated, given that the French facts that he cites are not entirely representative; for example, Jean-Yves Pollock (p.c.) finds acceptable:

(i) Jean semblait/avait semblé à Marie pouvoir faire l'affaire. ('J seemed/had seemed to M to-be-able to-do the trick')

For further discussion, see Bruening (2014: 713); it may be that the past (imperfect) tense of the finite verb or auxiliary in (i) favors full acceptability.

34. Despite the differences, there are also, as Anagnostopoulou (2016: 21) notes, "interpretational restrictions (definiteness, specificity, animacy) which are strikingly similar"; cf. Obenauer (1992) and Déprez (1998).

35. As noted by Morin (1985: 795).

This contrast between HCI in (53) and past participle agreement in (51) can be understood as follows. In (53) there is a clash between *te* and *elle*. This clash is due to the morpheme *-l-* that is part of *elle*. A clitic doubling relation cannot hold of two elements one of which contains third person *-l-* and the other of which is (first or) second person. Person agreement must hold with clitic doubling, as stated in (50). On the other hand, the past participle agreement morpheme *-e* in (51) contains no third person *-l-*; consequently, there is no person clash.³⁶

10. Missing persons

Of related importance is a restriction on HCI not yet mentioned, namely that the SCL found in HCI sentences must itself be third-person. Alongside the well-formed HCI example (48), or the following similar one:

(54) *Cela la gêne-t-elle?* ('that her bothers *t* she' = 'does that bother her?')

there is no comparable well-formed HCI example with a first or second person SCL. We can see this by starting with (52), which is an example of CI with a second person OCL *te*. If we then try to shift to HCI by making the SCL agree in person with that OCL, we reach:

(55) **Cela te gêne-(t-)tu?* ('that you bothers *t* you')

which is impossible. Similarly, alongside the well-formed HCI example:

(56) *Cela l'aurait-elle gênée?* ('that her would-have she bothered' = 'would that have bothered her?')

with a third person SCL *elle*, there is no parallel first person SCL example:³⁷

(57) **Cela m'aurait-je gêné?* ('that me would-have I bothered')

A question that arises is whether this person restriction on SCLs in HCI sentences is specific to HCI, or rather extends to CI. That is, can the SCL in CI sentences be first or second person? At first glance, there might seem to be well-formed CI sentences that do fit this description, e.g.:

36. Finite verb agreement shares with past participle agreement the absence of third person *-l-*; for a more detailed discussion, see Kayne (2003).

37. I have switched to a conditional tense because of restrictions on postverbal *je* discussed by Pollock (2006, note 43). In the first and second plural, the facts are the same:

(i) **Cela nous gêne-nous?*

(ii) **Cela vous gêne-vous?*

(58) Jean et moi avons-nous vu ce film? ('J and me have we seen that film')

However, (58) can alternatively be analyzed as left dislocation. Interference from left dislocation can be dampened (and a CI analysis more or less forced) by using sentences whose subject is quantified in a certain way.

In particular, Morin (1979, sect. 2.4) noted the contrast:

(59) Pourquoi lui seul a-t-il été prévenu? ('why him alone has *t* he been told')

(60) *Pourquoi toi seul as-tu été prévenu?³⁸ ('why you alone have *t* you been told')

In these examples, the subject DP contains *seul* ('alone', 'only') and there is a clear third person vs. non-third person contrast. The second person SCL *tu* in (60) is not possible. In a similar spirit, Pollock (2006: 622) used examples with a contrastive pronominal subject and found facts pointing in the same direction as Morin's:

(61) Quel livre lui a-t-il apporté? ('which book him has *t* he brought' = 'which book did HE bring?')

(62) *Quel livre moi ai-je apporté? ('which book me have *t* I brought')

Again, the non-third person (here first person) SCL *je* is not possible, with CI. Thus, both CI (as in (60) and (62)) and HCI (as in (55) and (57)) are impossible with a first or second person SCL.

The impossibility of CI and HCI with a first or second person SCL contrasts with first and second person examples of SCLI (in which the SCL is not doubling anything overt) as seen in:

(63) Aurais-je été prévenu? ('would-have I been told')

(64) As-tu été prévenu? ('have you been told')

(65) Avons-nous été prévenus? ('have we...')

(66) Avez-vous été prévenu(s)? ('have you...')

The well-formedness of (63)–(66) indicates clearly that CI and HCI are excluded from containing a first or second person SCL as a function of the clitic doubling that plays a central role in CI/HCI (vs. SCLI). The next question is why clitic doubling of the CI/HCI sort should be incompatible with first or second person.

38. Pollock (1983: 96) gives this example '*?' and a reviewer of the present paper says that it "does not sound so bad". It may be that some speakers can take *toi seul* to be dislocated or focussed, rather than being in subject position, in which case such examples would be examples of SCLI, not of CI; this would then account for the difference between the sometime acceptability of (60) and the strong and uniform rejection of (55), since the OCL *te* in (55) is not amenable to dislocation.

11. SCL *ce*

There is another restriction on SCLs in CI/HCI that is not found in SCLI. This restriction concerns the subject clitic *ce* of sentences like:³⁹

(67) *Ce n'est pas vrai.* ('that/it neg is not true')

which is related to the demonstrative *ce* of:⁴⁰

(68) *ce livre* ('that/this book')

The SCL *ce* of (67) is sometimes fully compatible with SCLI, as in:⁴¹

(69) *Est-ce vrai?* ('is that/it true')

(70) *Etait-ce vraiment comme cela?* ('was that/it really like that')

Surprisingly (at first glance), *ce* is not possible with CI.⁴² A relevant example, parallel to (60), is:⁴³

39. SCL *ce* is as a first approximation possible only with the verb 'be'; for details, see Kayne and Pollock (2010).

40. If the anti-homophony conjecture of Kayne (to appear) is correct, SCL *ce* and ordinary demonstrative *ce* must be exactly the same morpheme. The demonstrative character of SCL *ce* (suggested by Jean-Yves Pollock, p.c.) underlies its being unable to appear in core expletive-containing sentences like:

- (i) *Il est arrivé une lettre.* ('*il* is arrived a letter' = 'a letter has arrived')
- (ii) **C'est arrivé une lettre.*

with this contrast recalling:

- (iii) *There/*it has arrived a letter.*

and suggesting that (standard) English *it* might always be a (reduced) demonstrative.

41. Even when the verb is 'be', there are restrictions having to do with tense on *ce* in SCLI sentences that I take to be orthogonal to the present discussion.

42. There are no instances of *ce* with HCI, either, in part at least because there is (for reasons not yet discovered) no OCL *ce*:

- (i) **Jean ce sait.* ('J that/it knows')

This is so despite the fact that one finds, dialectally (cf. Bürge (1998)), sentences like:

- (ii) *Jean a ça vu.* ('J has that seen')

where *ça* has moved leftward in a way perhaps reminiscent of *tout* in standard French; on the movement of *tout*, Kayne (1975, Chapter 1), Pollock (1978), Starke (2001).

43. Other such examples are given in Kayne (1972: 83).

(71) * Pourquoi cela seul est-ce vrai? ('why that alone is that/it true')

Let me, then, pursue the idea that this restriction against *ce* in CI sentences is closely tied to the restriction against first and second person SCLs in CI (and HCI) sentences noted earlier in (55), (57), (60) and (62).

12. Demonstratives and first and second person pronouns

Kayne (2010a) proposed an account of the absence in English of a complementizer *this*. One component of that account was that *this* is invariably associated with a (perhaps unpronounced) first person morpheme.⁴⁴ Let me now complement that idea with the following:⁴⁵

(72) First and second person pronouns are invariably associated with demonstrative structure.

By 'demonstrative structure', I have in mind Leu's (2007) proposal⁴⁶ that demonstratives are phrasal and contain the definite article as a subpart. Combined with (72), this yields (73):

(73) First and second person pronouns are invariably accompanied by a (usually silent) definite article.

This is illustrated in:

(74) THE you/me/us PERSON(S)

where capitalization indicates silence.⁴⁷

We are now in a position to return to the restriction against demonstrative-like SCL *ce* in CI sentences discussed in the previous section and to the closely related restriction against first and second person SCLs in CI (and HCI) sentences noted earlier in (55), (57), (60) and (62). If the proposal in (72) is correct, these two restrictions boil down to one:

44. Cf. Leu (2007, note 2) and references cited there.

45. Cf. in part Jayaseelan and Hariprasad (2001).

46. Which has various antecedents. See Leu (2007, Introduction).

47. On PERSON, cf. Kayne (2005a, Appendix). The definite article can be pronounced in:

(i) That's not the you that everybody used to love.

The text proposal differs from Postal (1966), Ritter (1995) and Déchaine and Wiltschko (2002), who take first and second person pronouns to be determiners.

- (75) CI and HCI are incompatible with SCLs associated with phrasal demonstrative structure.

If we now ask why (75) should hold, a possible (beginning of an) answer is that the postverbal agreeing SCLs of CI and HCI must not be too complex.⁴⁸ SCLs associated with phrasal demonstrative structure (*ce* and, by extension from (74), first and second person SCLs) would, then, be too complex to be compatible with HCI or CI,⁴⁹ while third person SCLs would be less complex.⁵⁰

13. Other types of clitic doubling

CI and HCI are subtypes of clitic doubling. French has another subtype involving OCLs, seen in:

- (76) *Ils la voient elle.* ('they her see her')

in which there is a contrastive interpretation and intonation, such that (76) is distinct from right dislocation. The kind of OCL clitic doubling illustrated in (76) (which in French requires that the doubled phrase in argument position be a

48. A reviewer notes an intriguing similarity here to colloquial Central-Oriental Catalan, which allows clitic reduplication of the sort seen in:

- (i) *Ho vol fer-ho* ('it (s)he-wants to-do it')

This kind of reduplication is marginally available with *l*-clitics, but not at all with the clitics picked out by (75) (as opposed to Bellinzonese; cf. Cattaneo 2009, sect. 6.7).

49. Despite first and second person plural SCLs being able to co-occur, in SCLI, with distinctive suffixal agreement morphemes, as in:

- (i) *Partez-vous?*
 (ii) *Partons-nous?*

with *part-* ('leave') the verbal root, *vous* ('you') and *nous* ('we') the SCLs, and *-ez* and *-ons* the corresponding agreement morphemes.

50. Possibly, (postverbal) third person SCLs are not phrasal; if so, then since third person SCLs show gender and number morphology (on number, see Kayne and Pollock 2014, sect. 7), they would have to have been 'put together' by head movement. Alternatively, it is the deictic subpart of demonstratives (and first and second person pronouns) that makes the difference.

For an argument, differing from Postal (1966), to the effect that pronouns (in Russian) are nouns, see Franks and Pereltsvaig (2004). On the varying structural complexity of pronouns, see Cardinaletti and Starke (1999) and Déchaine and Wiltschko (2002). On French *celui* ('that him') as an instance of (non-agreeing) determiner + third person (strong, non-clitic) pronoun, see Kayne (2010a, sect. 10).

pronoun) differs from CI and HCI (as does Spanish clitic doubling) in being compatible with first and second person pronouns, e.g.:

(77) Ils te voient toi. ('they you see you')

French allows this kind of non-dislocation contrastive doubling with subject pronouns, too,⁵¹ as in:

(78) Elle partira elle. ('she will-leave she/her')

and again allows it with first and second person pronouns, e.g.:

(79) Je partirai moi. ('I will-leave I/me')

The question now is why (79), with a first person SCL, should differ in acceptability from the unacceptable CI example (62), repeated here:

(80) *Quel livre moi ai-je apporté? ('which book me have *t* I brought')

as well as from the sharply unacceptable HCI example (57), also repeated:

(81) *Cela m'aurait-je gêné? ('that me would-have I bothered')

(and similarly for other first and second person SCLs).

A possible answer goes as follows. CI and HCI involve a complex DP analysis of the sort indicated in Section 2, in which the SCL and its double both start out within one DP. A complex DP of that sort is not compatible with SCLs associated with phrasal demonstrative structure, i.e. with *ce* or with first or second person SCLs, whence the unacceptability of (80) and (81). On the other hand, (79) does not involve such a complex DP structure, and is therefore possible. (By extension, (77) will not involve a complex DP structure, either.)

51. Cf. Ronat (1979).

As for the correct derivation of (77) and (79) (and, possibly, (76) and (78)),⁵² one might consider extending to those sentences the kind of analysis envisaged in Kayne (1994, sect. 8.3) for:⁵³

(82) He's real smart, John.

in which this right-dislocation derives from the biclausal:

(83) He's real smart, John is.

Transposed to (77) and (79), this would amount to taking them to be something like:⁵⁴

(84) ils te voient ILS VOIENT toi

and:

(85) je partirai moi PARTIRAI

with capitalization again indicating silence.

The proposal indicated in (85) can be maintained even though *moi* is not normally a possible subject by itself:

(86) *Moi partirai.

52. Gatti (1989/90, 195n) pointed out for Trentino a difference between 1st/2nd and 3rd person with respect to clitic doubling:

(i) I me vede mi ('they me see me')

(ii) I te vede ti ('they you see you')

(iii)*? I la vede ela ('they her see her')

Non-clitic *mi/ti* can cooccur in Trentino with clitic *me/te*, but non-clitic *ela* cannot cooccur with clitic *la*. Cf. Burzio (1989) on Piedmontese. This contrast seems to hold for Paduan, too (Paola Benincà, p.c.); cf. Benincà (1983, note 8). On the other hand, it seems to be absent from the dialects studied by Nicoli (1983: 144, 359), Pellicardi (1977: 93), Vassere (1993: 97, 102), and Spiess (1976: 209). Future work should individuate the parameter(s) underlying this difference.

53. Cf. also Ott (2014). This kind of analysis must then not be available to CI/HCI. Though it must be available to contrastive doubling even with *ne...que* added, as in:

(i) Cela ne te plaît qu'à toi. ('that neg you pleases than to you' = 'that pleases only you')

54. And similarly for Spanish clitic doubling, at least with first and second person OCLs. Possibly, there's a link here to sentences like the following, in some Italian:

(i) È andato a Parigi è andato. ('he-is gone to P he-is gone')

For discussion, see Gulli (2003).

This is so (and similarly for (84)), since *moi* can be a subject by itself in gapping examples like:

- (87) Jean aime la physique et moi la chimie. ('J likes the physics and I the chemistry')

in which there must be a silent verb in the second part of the sentence, as, then, in (85).⁵⁵

14. The sensitivity of silent pronouns to person

The difference in structure suggested in (72) between first and second person pronouns, which are associated with phrasal demonstrative structure, and third person pronouns, which are not, may find additional support in the behavior of certain silent pronouns, in a way that can be seen in French. For example, French allows:⁵⁶

- (88) Tous chantaient. ('all were-singing-3PL')

in which there must certainly be a silent third person pronoun within the subject DP. Of note is the fact that sentences like (88) are limited to third person subjects, as shown in:

- (89) *Tous chantiez. ('all were-singing-2PL')

- (90) *Tous chantions. ('all were-singing-1PL')

All of (88)–(90) have distinctive verbal agreement endings. They can nonetheless be distinguished in acceptability if we take French to allow a silent third person pronoun as part of the subject in (88), but to disallow comparable silent first or second person pronouns.

English shows similar behavior, as seen in:

- (91) Both/all five were behaving themselves yesterday morning.

There must again be a silent third person pronoun here, accompanying *both* and *all five*, within the subject DP:

- (92) both/all five THEM were...

55. If Johnson (2009) is correct and extendable to gapping in comparatives, the silent verb in gapping would come about as the result of across-the-board movement. Possibly, (79)/(85) contains a silent JE, too.

56. For additional details, see Kayne (2001).

As in French, this silent pronoun cannot be first or second person:

(93) *Both/all five were behaving ourselves/yourselves yesterday morning.

Similarly, we have:

(94) Five/most/not very many were behaving themselves/ *ourselves/*yourselves yesterday morning.

again with a silent (OF) THEM that has no first or second person counterpart.

Italian shows similar behavior in the particular case of:

(95) *Quattro sono venuti.* ('four are-3PL. come' = 'four of them have come')

(96) **Quattro siete venuti.* ('four are-2PL. come')

On the other hand, Italian allows sentences of the sort seen in (89)/(90), for example:

(97) *Tutti siamo felici.* ('all are(1PL.) happy')

Since Italian is a robust null subject language, (97) is possible with an analysis in which *tutti* is not in subject position. This amounts to saying that the acceptability of (97) doesn't depend on the presence of *tutti*, as shown in fact by:

(98) *Siamo felici.*⁵⁷

That French (89) or (90) is not possible now reduces to the fact that French, not being a null subject language, does not allow:

(99) **Chantions.* ('were singing')

In the same way English (93) is impossible exactly as is:

(100) **Were behaving ourselves/yourselves yesterday morning.*

The fact that within Italian (96) contrasts with (97) can now be related to the fact that *tutti* is a possible floating/stranded quantifier, just as English *all* is, whereas *quattro* is not, just as English *four* is not:

(101) *They are all here.*

(102) **They are four here.*

The facts of (88)–(102) taken together support the following cautious statement:

(103) Silent pronouns are sometimes limited to third person.

57. Here the true subject may be *pro* or it may be the agreement suffix *-m(o)*; cf. Taraldsen (1992).

The caution is warranted by the fact that Italian itself allows a silent first person pronoun in:

(104) *Vogliono che parta.* ('they-want that leave')

In this example, the embedded verb *parta* is present subjunctive and is notable in that its *-a* suffix is in all probability not an agreement morpheme (but rather a theme vowel). Despite the lack of any overt first person agreement morpheme, (104) can have the subject of *parta* interpreted as first person (singular). This leads to the conclusion that (104) can contain a silent first person singular pronoun. (There is also a possible third singular interpretation.)

One factor relevant to (104) vs. (88)–(96) is that in the latter set of cases, the silent pronoun is a subpart of the subject, not the whole subject, which suggests that a canonical pro-drop configuration provides an extra licensing possibility.⁵⁸ Setting that aside, let me propose that the limitation to third person in (88)–(96) is to be understood in terms of (72), i.e. in terms of the idea that first and second person pronouns are associated with a demonstrative structure, whose silence in contexts like those of (88)–(96) can evidently not be licensed in the way that the silence of less (or differently) complex third person pronouns can be.

15. French *on*

The French SCL *on* (which in certain other cases can correspond to English generic subject *one*) can pair with *nous* ('we/us') in sentences like:

(105) *Nous, on va à Paris.* ('us on go to P' = 'we're going to P')

(106) *On va à Paris, nous.*

with a first person plural interpretation.⁵⁹ Yet alongside the CI example:

(107) *Cela nous gêne-t-il?* ('that us bothers it' = 'does that bother us?')

there is no HCI-like:

(108) **Cela nous gêne-t-on?*

58. Somewhat similarly, all instances of PRO are apparently indifferent to person; the silent subject of at least non-agreeing imperatives might be PRO, thinking of Ross (1970). French *voici/voilà* might, exceptionally (for French), have a(n obligatorily) silent non-PRO second person subject; cf. Morin (1985: 817).

59. For relevant discussion concerning a comparable property of Italian *si*, see Cinque (1988).

This is so even though *on* is compatible with SCLI:

- (109) A-t-on tous fait la même erreur? ('has on all made the same mistake' = 'have we all...?')

The incompatibility of *on* with HCI can also be seen using so-called middle sentences like:

- (110) Cela se lit facilement. ('this book *se* reads easily')

which is very close in interpretation to:

- (111) On lit cela facilement.

Of interest here is the fact that middles are compatible with CI, with SCL = *il*:

- (112) Cela se lit-il facilement?

but not with HCI:

- (113) *Cela se lit-on facilement?

i.e. HCI-type doubling of *se* by *on* is prohibited.

Similarly, although the following two sentences are close in interpretation:

- (114) Quelqu'un vous attend. ('someone you awaits' = 'someone awaits you')

- (115) On vous attend.

we have, with HCI:

- (116) Quelqu'un vous attend-il?

but not:

- (117) *Quelqu'un vous attend-on?

The incompatibility of *on* with HCI seen in (108), (113) and (117) suggests that *on* should be grouped with *ce*, *je* and *tu* and that as with those SCLs we should attribute to *on* demonstrative structure.⁶⁰ Grouping *on* with first and second person pronouns (despite its triggering the same verb agreement as third singular)⁶¹ is

60. If so, that would lead in a different direction from Kayne (1972: 95)'s taking *nous* and *on* to be part of one DP-like phrase.

61. And despite its differing with respect to coordination (*on* can be dropped from a second conjunct less readily than *je*, *tu*; cf. Kayne 1975, Chapter 2, notes 37, 40, Sportiche 1999, sect. 5.2), in a way that may be linkable to the fact that Italian third person counterparts to French *on* sentences cannot be without *si*.

supported by the fact that no subtype of *on* ever varies in form for gender, just as first and second person pronouns never do, in Romance.⁶²

This grouping of *on* with first and second person pronouns is also indirectly supported by the parallelism between French *on* and Italian impersonal *si* discussed by Cinque (1988, sect. 3.5). This is so if *on* is a nominative counterpart of French *se*, as suggested by Togeby (1982: 428), if all instances of *se* are the same element,⁶³ and if, as in Kayne (2003), *se* and *si* are themselves to be grouped with first and second person singular pronouns.⁶⁴

16. Conclusion

French Hyper-Complex Inversion (HCI) is an instance of clitic doubling that is subject to a person restriction not found with more familiar cases of clitic doubling. This restriction is argued to result from an incompatibility between the postverbal subject clitic (SCL) of HCI and the demonstrative structure associated with first and second person pronouns. That demonstrative structure also plays a role in asymmetries that hold concerning the possible silence, in certain cases, of third person pronouns, but not first or second person pronouns.

HCI shares with past participle agreement the property that it is incompatible with an unmoved lexical direct object, in a way that presents a challenge to Agree, if Agree is taken to be available even in the absence of movement.

The SCL of HCI sentences can sometimes climb out of an embedded infinitive, in a way related to the Comp-area character of its landing site.

62. This is completely clear for 1st/2nd singular (cf. Kayne 2003). Spanish 1st and 2nd plural *nosotros*, *vosotros* have feminine counterparts *nosotras*, *vosotras*; rather than taking them to be exceptions, as in Dobrovie-Sorin and Giurgea (2011: 134), I take them to have a non-agreeing pronominal subpart *nos*, *vos* that is followed by an agreeing non-pronominal *otras*, the feminine plural form of *otro* ('other').

63. Necessarily so if Kayne's (to appear, (17)) anti-homophony conjecture (cf. Embick 2003: 146, 156, for an earlier, more flexible version) is correct; cf. also Leu (2017).

64. Which would suggest, from the present perspective, that all instances of *se/si* are associated with demonstrative structure, with the possibility then arising that all instances of *se/si* have something in common with expletive *there*, in particular if expletive *there* originates DP-internally, as in Kayne (2008). More specifically, it may be that *se/si* is to *where* as 1st and 2nd person singular pronouns are to *here* and *there*; on the status of *-r-* in these, see Noonan (2017).

References

- Anagnostopoulou, Elena. 2016. "Clitic Doubling and Object Agreement." In *Proceedings of the VII Nereus International Workshop Clitic Doubling and other issues of the syntax/semantic interface in Romance DPs*, ed. by Susann Fischer, and Mario Navarro, 11–42. Konstanz: Fachbereich Sprachwissenschaft der Universität Konstanz.
- Baker, Mark. 2008. *The syntax of agreement and concord*. Cambridge: Cambridge University Press. <https://doi.org/10.1017/CBO9780511619830>
- Belletti, Adriana. 2006. "(Past) Participle Agreement." In *The Blackwell Companion to Syntax, Volume III*, ed. by Martin Everaert, and Henk van Riemsdijk, 493–521. Malden, MA: Blackwell. <https://doi.org/10.1002/9780470996591.ch48>
- Benincà, Paola. 1983. "Il clitico *a* nel dialetto padovano." In *Scritti linguistici in onore di G.B. Pellegrini*, 25–35. Pisa: Pacini. [reprinted in Benincà (1994)].
- Benincà, Paola. 1994. *La variazione sintattica. Studi di dialettologia romanza*. Bologna: Il Mulino.
- Boneh, Nora, and Ivy Sichel. 2010. "Deconstructing Possession." *Natural Language and Linguistic Theory* 28: 1–40. <https://doi.org/10.1007/s11049-009-9087-z>
- Bonet, Eulalia. 1991. "Morphology After Syntax: Pronominal Clitics In Romance." Ph.D. dissertation, MIT.
- Bruening, Benjamin. 2014. "Defects of Defective Intervention." *Linguistic Inquiry* 45: 707–719. https://doi.org/10.1162/LING_a_00171
- Bürgi, Anne. 1998. "Le pronom *ça* en français vaudois, description et analyse." Mémoire de Maîtrise, University of Sherbrook.
- Burzio, Luigi. 1989. "Work in Progress." Ms., Harvard University.
- Butt, John B., and Carmen Benjamin. 1988. *A New Reference Grammar of Modern Spanish*. London: Edward Arnold.
- Caha, Pavel. 2010. "The German Locative-Directional Alternation. A Peeling Account." *Journal of Comparative Germanic Linguistics* 13: 179–223. <https://doi.org/10.1007/s10828-010-9039-3>
- Cardinaletti, Anna, and Michal Starke. 1999. "The Typology of Structural Deficiency: A Case Study of the Three Classes of Pronouns." In *Clitics in the Languages of Europe*, ed. by Henk van Riemsdijk, 145–233. Berlin: Mouton de Gruyter.
- Cattaneo, Andrea. 2009. "It Is All About Clitics: The Case of a Northern Italian Dialect Like Bellinzonese." Ph.D. dissertation, New York University.
- Chomsky, Noam. 2000. "Minimalist Inquiries: The Framework," in *Step by Step: Essays in Minimalist Syntax in Honor of Howard Lasnik*, ed. by Roger Martin, David Michaels, and J. Uriagereka, 89–155. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2001. "Derivation by Phase." In *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2008. "On Phases." In *Foundational Issues in Linguistic Theory. Essays in Honor of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 133–166. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9780262062787.003.0007>
- Cinque, Guglielmo. 1988. "On *Si* Constructions and the Theory of Arb." *Linguistic Inquiry* 19:521–581.
- Cinque, Guglielmo. 1990. *Types of A'-Dependencies*. Cambridge, MA: MIT Press.

- Cinque, Guglielmo. 1995. *Italian Syntax and Universal Grammar*. Cambridge: Cambridge University Press.
- Collins, Chris. 2005. "A Smuggling Approach to the Passive in English." *Syntax* 8:81–120. <https://doi.org/10.1111/j.1467-9612.2005.00076.x>
- Déchainé, Rose-Marie, and Martina Wiltschko. 2002. "Decomposing Pronouns." *Linguistic Inquiry* 33: 409–442. <https://doi.org/10.1162/002438902760168554>
- Déprez, Viviane. 1998. "Semantic Effects of Agreement: The Case of French Past Participle Agreement." *Probus* 10: 1–65. <https://doi.org/10.1515/prbs.1998.10.1.1>
- Dobrovie-Sorin, Carmen, and Ion Giurgea. 2011. "Pronominal Possessors and Feature Uniqueness." *Language* 87: 126–157. <https://doi.org/10.1353/lan.2011.0020>
- Embick, David. 2003. "Locality, Listedness, and Morphological Identity." *Studia Linguistica* 57: 143–169. <https://doi.org/10.1111/j.0039-3193.2003.00102.x>
- Franks, Steven, and Asya Pereltsvaig. 2004. "Functional Categories in the Nominal Domain." In *Proceedings of Formal Approaches to Slavic Linguistics 12. The Ottawa Meeting 2003*, ed. by O. Arnaudova et al., 109–128. Ann Arbor: University of Michigan Press.
- Gatti, Tiziana. 1989/1990. "Confronto tra fenomeni sintattici nell'italiano e nel dialetto trentino: participio passato, accordo e ausiliari." Tesi di laurea, University of Trento.
- Gruber, Bettina. 2008. "Complementiser Agreement – New Evidence from the Upper Austrian Variant of Gmunden." Master's thesis, University of Vienna.
- Gulli, Antonino. 2003. "Reduplication in Syntax." Ph.D. dissertation, Graduate Center, City University of New York.
- Jayaseelan, K. A., and M. Hariprasad. 2001. "Deixis in Pronouns and Noun Phrases." *Linguistic Analysis* 31: 132–149.
- Johnson, Kyle. 2009. "Gapping Is Not (VP-) Ellipsis." *Linguistic Inquiry* 40: 289–328. <https://doi.org/10.1162/ling.2009.40.2.289>
- Kayne, Richard S. 1972. "Subject Inversion in French Interrogatives." In *Generative Studies in Romance Languages*, ed. by Jean Casagrande, and Bohdan Saciuk, 70–126. Rowley: Newbury House.
- Kayne, Richard S. 1975. *French Syntax. The Transformational Cycle*. Cambridge, MA: MIT Press.
- Kayne, Richard S. 1981. "Binding, quantifiers, clitics and control." In *Binding and filtering*, ed. by Frank Heny, 191–211. London: Croom Helm. [reprinted in Kayne (1984)].
- Kayne, Richard S. 1983. "Chains, Categories External to S and French Complex Inversion." *Natural Language and Linguistic Theory* 1: 107–139. [reprinted in Kayne (1984)]. <https://doi.org/10.1007/BF00210378>
- Kayne, Richard S. 1984. *Connectedness and Binary Branching*. Dordrecht: Foris.
- Kayne, Richard S. 1985. "L'accord du participe passé en français et en italien." *Modèles Linguistiques* VII: 73–90. [English translation in Kayne (2000), Chapter 2].
- Kayne, Richard S. 1989a. "Null Subjects and Clitic Climbing." In *The Null Subject Parameter*, ed. by Osvaldo Jaeggli, and Kenneth Safir, 239–261. Dordrecht: Reidel. [reprinted in Kayne (2000)]. https://doi.org/10.1007/978-94-009-2540-3_8
- Kayne, Richard S. 1989b. "Facets of Romance Past Participle Agreement." In *Dialect Variation and the Theory of Grammar*, ed. by Paola Benincà, 85–103. Dordrecht: Foris. [reprinted in Kayne (2000)].
- Kayne, Richard S. 1991. "Romance Clitics, Verb Movement and PRO." *Linguistic Inquiry* 22: 647–686. [reprinted in Kayne (2000)].
- Kayne, Richard S. 1993. "Toward a Modular Theory of Auxiliary Selection." *Studia Linguistica* 47: 3–31. [reprinted in Kayne (2000)]. <https://doi.org/10.1111/j.1467-9582.1993.tb00837.x>

- Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Kayne, Richard S. 1998. "Overt vs. Covert Movement." *Syntax* 1: 128–191. [reprinted in Kayne (2000)]. <https://doi.org/10.1111/1467-9612.00006>
- Kayne, Richard S. 2000. *Parameters and Universals*. New York: Oxford University Press.
- Kayne, Richard S. 2001. "A Note on Clitic Doubling in French." In *Current Studies in Italian Syntax. Essays offered to Lorenzo Renzi*, ed. by Guglielmo Cinque, and Giampaolo Salvi, 189–212. Amsterdam: North-Holland. [also in Kayne (2000)].
- Kayne, Richard S. 2002. "Pronouns and Their Antecedents." In *Derivation and Explanation in the Minimalist Program*, ed. by Samuel Epstein, and David Seely, 133–166. Malden, MA: Blackwell. [reprinted in Kayne (2005b)]. <https://doi.org/10.1002/9780470755662.ch7>
- Kayne, Richard S. 2003. "Person Morphemes and Reflexives in Italian, French and Related Languages." In *The Syntax of Italian Dialects*, ed. by Christina Tortora, 102–136. New York: Oxford University Press. [also in Kayne (2000)].
- Kayne, Richard S. 2005a. "On the Syntax of Quantity in English." In Kayne (2005b). Also in *Linguistic Theory and South-Asian Languages. Essays in Honour of K.A. Jayaseelan*, ed. by Josef Bayer, Tanmoy Bhattacharya, and M. T. Hany Babu, 73–105. Amsterdam: John Benjamins.
- Kayne, Richard S. 2005b. *Movement and Silence*. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195179163.001.0001>
- Kayne, Richard S. 2008. "Expletives, Datives, and the Tension between Morphology and Syntax." In *The Limits of Syntactic Variation*, ed. by Theresa Biberauer, 175–217. Amsterdam: John Benjamins. [reprinted in Kayne (2010b)]. <https://doi.org/10.1075/la.132.07kay>
- Kayne, Richard S. 2009. "Some Silent First Person Plurals." In *Merging Features. Computation, Interpretation, and Acquisition*, ed. by José M. Brucart, Anna Gavarró, and Jaume Solà, 276–292. New York: Oxford University Press. [reprinted in Kayne (2010b)]. <https://doi.org/10.1093/acprof:oso/9780199553266.003.0015>
- Kayne, Richard S. 2010a. "Why Isn't This a Complementizer?" In Kayne (2010b). Also in *Functional Structure from Top to Toe: A Festschrift for Tarald Taraldsen*, ed. by Peter Svenonius, 188–231. New York: Oxford University Press, 2014.
- Kayne, Richard S. 2010b. *Comparisons and Contrasts*. New York: Oxford University Press.
- Kayne, Richard S. To appear. "The Unicity of There and the Definiteness Effect." (lingbuzz/002858).
- Kayne, Richard S., and Jean-Yves Pollock. 2001. "New Thoughts on Stylistic Inversion." In *Inversion in Romance*, ed. by Aafke Hulk, and Jean-Yves Pollock, 107–162. New York: Oxford University Press. [reprinted in Kayne (2005)].
- Kayne, Richard S., and Jean-Yves Pollock. 2010. "Notes on French and English Demonstratives." in *Structure Preserved: Studies in Syntax for Jan Koster*, ed. by Jan-Wouter Zwart, and Mark de Vries, 215–228. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.164.24kay>
- Kayne, Richard S., and Jean-Yves Pollock. 2012. "Toward an Analysis of French Hyper-Complex Inversion." In *Functional Heads*, ed. by Laura Brugè, Anna Cardinaletti, Giuliana Giusti, Nicola Munaro, and Cecilia Poletto, 150–167. New York: Oxford University Press. [also in Kayne (2010b)]. <https://doi.org/10.1093/acprof:oso/9780199746736.003.0012>
- Kayne, Richard S., and Jean-Yves Pollock. 2014. "Locality and Agreement in French Hyper-Complex Inversion." In *Locality*, ed. by Enoch Olade Aboh, Maria Teresa Guasti, and Ian Roberts, 32–57. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199945269.003.0002>

- Koopman, Hilda. 2003. "The locality of agreement and the structure of the DP in Maasai." In *The Role of Agreement in Natural Language: TLS 5 Proceedings*, ed. by William Earl Griffin, 206–227. Texas Linguistic Forum, 53.
- Koopman, Hilda. 2005a. "Korean (and Japanese) Morphology from a Syntactic Perspective." *Linguistic Inquiry* 36: 601–633. <https://doi.org/10.1162/002438905774464359>
- Koopman, Hilda. 2005b. "On the Parallelism of DPs and Clauses. Evidence from Kisongo Maasai." In *Verb First, On the Syntax of Verb-initial Languages*, ed. by Andrew Carnie, Heidi Harley, and Sheila Ann Dooley, 281–302. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.73.19koo>
- Koopman, Hilda, and Anna Szabolcsi. 2000. *Verbal Complexes*. Cambridge, MA: MIT Press.
- Leu, Thomas. 2007. "These HERE demonstratives." In *Proceedings of the 30th Annual Penn Linguistics Colloquium*, ed. by Tatjana Scheffler, Joshua Tauberer, Aviad Eilam, and Laia Mayol, 141–154. Philadelphia: Penn Working Papers in Linguistics.
- Leu, Thomas. 2017. "Ein is Ein and that is that: A note on anti-homophony and meta-morphology." In *The Structure of Words at the Interfaces*, ed. by Heather Newell, Máire Noonan, Glyne Piggott, and Lisa deMena Travis. New York: Oxford University Press.
- Longobardi, Giuseppe. 1985. "Connectedness, scope, and c-command." *Linguistic Inquiry* 16: 163–192.
- Morin, Yves-Charles. 1979. "There is no Inversion of Subject Clitics in Modern French." Ms., University of Montreal.
- Morin, Yves-Charles. 1985. "On the Two French Subjectless Verbs *Voici* and *Voilà*." *Language* 61: 777–818. <https://doi.org/10.2307/414490>
- Napoli, Donna Jo. 1981. "Subject pronouns: the pronominal system of Italian vs. French." *Papers from the 17th Regional Meeting of the Chicago Linguistic Society*, ed. by Roberta A. Hendrick, Carrie S. Masek, Mary Frances Miller, 249–276. Chicago: Chicago Linguistic Society.
- Navarro, Mario, and Mareike Neuhaus. 2016. "Clitic Doubling restrictions in Leísta Spanish." In *Proceedings of the VII Nereus International Workshop Clitic Doubling and other issues of the syntax/semantic interface in Romance DPs*, ed. by Susann Fischer, and Mario Navarro, 79–89. Konstanz: Fachbereich Sprachwissenschaft der Universität Konstanz.
- Nicoli, Franco. 1983. *Grammatica milanese*. Busto Arsizio: Bramante.
- Noonan, Máire. 2017. "Dutch and German *r*-pronouns and P-stranding." In *The Structure of Words at the Interfaces*, ed. by Heather Newell, Máire Noonan, Glyne Piggott, and Lisa deMena Travis, 209–239. New York: Oxford University Press.
- Obenauer, Hans-Georg. 1992. "L'interprétation des structures *wh* et l'accord du participe passé." In *Structures de la phrase et théorie du Liage*, ed. by Hans-Georg Obenauer, and Anne Zribi Hertz, 169–193. Saint-Denis: Presses Universitaires de Vincennes.
- Ordóñez, Francisco, and Estela Treviño. 1999. "Left Dislocated Subjects and the Pro-drop Parameter: A Case Study of Spanish." *Lingua* 107: 39–68. [https://doi.org/10.1016/S0024-3841\(98\)00020-5](https://doi.org/10.1016/S0024-3841(98)00020-5)
- Ott, Dennis. 2014. "An Ellipsis Approach to Contrastive Left-Dislocation." *Linguistic Inquiry* 45: 269–303. https://doi.org/10.1162/LING_a_00155
- Pellicciardi, Ferdinando. 1977. *Grammatica del dialetto romagnolo*. Ravenna: Longo.
- Penello, Nicoletta. 2003. "Capitolo di morfologia e sintassi del dialetto di Carmignano di Brenta." Tesi di dottorato, Università di Padova.
- Penello, Nicoletta. 2007. "On Double Clitics in Interrogatives in a Northern Italian Dialect." *Nordlyd* 34: 201–217. <https://doi.org/10.7557/12.95>

- Poletto, Cecilia. 2000. *The Higher Functional Field in the Northern Italian Dialects*. New York: Oxford University Press.
- Pollock, Jean-Yves. 1978. "Trace Theory and French Syntax." In *Recent Transformational Studies in European Languages*, ed. by Samuel J. Keyser, 65–112. Cambridge, MA: MIT Press.
- Pollock, Jean-Yves. 1983. "Sur quelques propriétés des phrases copulatives en français." *Langue Française* 58: 89–125. <https://doi.org/10.3406/lfr.1983.6416>
- Pollock, Jean-Yves. 1985. "On Case and the Syntax of Infinitives in French." In *Grammatical Representation*, ed. by Jacqueline Guéron, Hans-Georg Obenauer, and Jean-Yves Pollock, 293–326. Dordrecht: Foris.
- Pollock, Jean-Yves. 1989. "Verb Movement, Universal Grammar, and the Structure of IP." *Linguistic Inquiry* 20: 365–424.
- Pollock, Jean-Yves. 2006. "Subject Clitics and Complex Inversion." In *The Blackwell Companion to Syntax, Volume IV*, ed. by Martin Everaert, and Henk van Riemsdijk, 601–659. Malden, MA: Blackwell. <https://doi.org/10.1002/9780470996591.ch67>
- Postal, Paul M. 1966. "On So-Called 'Pronouns' in English." In *Report of the Seventeenth Annual Roundtable Meeting on Linguistics and Language Studies*, ed. by Francis P. Dineen, 177–206. Washington, D.C.: Georgetown University Press. [reprinted in *Modern Studies in English*, ed. by David A. Reibel, and Sanford A. Schane, 201–224. Englewood Cliffs, NJ: Prentice-Hall, 1969.]
- Preminger, Omer. 2009. "Breaking Agreements: Distinguishing Agreement and Clitic Doubling by Their Failures." *Linguistic Inquiry* 40: 619–666. <https://doi.org/10.1162/ling.2009.40.4.619>
- Ritter, Elizabeth. 1995. "On the Syntactic Category of Pronouns and Agreement." *Natural Language and Linguistic Theory* 13: 405–443. <https://doi.org/10.1007/BF00992737>
- Rizzi, Luigi. 1997. "The fine structure of the left periphery." In *Elements of Grammar*, ed. by Liliane Haegeman, 281–337. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5420-8_7
- Roberts, Ian. 1993. "The Nature of Subject Clitics in Franco-Provençal Valdôtain." In *Syntactic Theory and the Dialects of Italy*, ed. by Adriana Belletti, 319–353. Torino: Rosenberg & Sellier.
- Roberts, Ian. 2010. *Agreement and Head Movement. Clitics, Incorporation, and Defective Goals*. Cambridge, MA: MIT Press. <https://doi.org/10.7551/mitpress/9780262014304.001.0001>
- Ronat, Mitsou. 1979. "Pronoms topiques et pronoms distinctifs." *Langue Française* 44: 106–128. <https://doi.org/10.3406/lfr.1979.6172>
- Ross, John Robert. 1970. "On Declarative Sentences." In *Readings in English Transformational Grammar*, ed. by Roderick A. Jacobs, and Peter S. Rosenbaum, 222–272. Waltham, MA: Ginn.
- Rouveret, Alain, and Jean-Roger Vergnaud. 1980. "Specifying Reference to the Subject: French Causatives and Conditions on Representations." *Linguistic Inquiry* 11: 97–202.
- Schoorlemmer, Erik. 2006. "Double Agreement in Complex Inversion." In *Romance Languages and Linguistic Theory 2004. Selected papers from 'Going Romance', Leiden, 9–11 December 2004*, ed. by Jenny Doetjes, and Paz González, 275–295. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.278.13sch>
- Shlonsky, Ur. 1994. "Agreement in Comp." *The Linguistic Review* 11: 351–375. <https://doi.org/10.1515/tlir.1994.11.3-4.351>

- Spiess, Federico. 1976. "Di un'innovazione morfologica nel sistema dei pronomi personali oggetto del dialetto della Collina d'Oro." In *Problemi di morfosintassi dialettale*, 203–212. Pisa: Pacini.
- Sportiche, Dominique. 1999. "Subject Clitics in French and Romance. Complex Inversion and Clitic Doubling." In *Beyond Principles and Parameters*, ed. Kyle Johnson, and Ian Roberts, 189–221. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-4822-1_8
- Starke, Michal. 2001. "Move Dissolves into Merge: A Theory of Locality." Ph.D. dissertation, University of Geneva.
- Szabolcsi, Anna. 1983. "The Possessor that Ran Away from Home." *The Linguistic Review* 3: 89–102. <https://doi.org/10.1515/tlir.1983.3.1.89>
- Szabolcsi, Anna. 1994. "The Noun Phrase." In *Syntax and Semantics 27. The Syntactic Structure of Hungarian*, ed. by Ferenc Kiefer, and Katalin É. Kiss, 179–274. San Diego: Academic Press.
- Taraldsen, K. Tarald. 1992. "Agreement as Pronoun Incorporation." Paper presented at the 15th GLOW Colloquium. *GLOW Newsletter* 28: 50–51.
- Togeby, Knud. 1982. *Grammaire Française, Volume I: Le Nom*. Copenhagen: Akademisk Forlag.
- Torrego, Esther. 1996. "On Quantifier Float in Control Clauses." *Linguistic Inquiry* 27: 111–126.
- Uriagereka, Juan. 1995. "Aspects of the Syntax of Clitic Placement in Western Romance." *Linguistic Inquiry* 26: 79–123.
- Vassere, Stefano. 1993. *Sintassi formale e dialettologia. I pronomi clitici nel luganese*. Milan: FrancoAngeli.
- Zagona, Karen. 2002. *The syntax of Spanish*. Cambridge: Cambridge University Press.

Licensing conditions on null generic subjects in Spanish

Matthew L. Maddox

University of Illinois at Urbana-Champaign & Southeast Community College-Nebraska

Holmberg's (2005, 2010) Null Generic Subject Generalization (NGSG) states that only partial null subject languages allow null generic third-singular subjects while consistent null subject languages (cNSLs) like Spanish only have referential *pro* due to a D-feature in T. Novel data from Spanish contradicts this. Following Frascarelli (2007), I analyze the licensing of null generic subjects as topic-identification with generic *uno*. The Special Morphology Condition on the NGSG, which stipulates that generic *pro* is allowed in cNSLs when licensed by special morphology, cannot account for *uno*. Thus, we may dispense with the postulation of a D-feature in T in cNSLs. I propose a revised set of licensing conditions on *pro* in null-subject languages.

Keywords: generic *pro*, generic *uno*, null subjects, impersonal *se*, topic-identification

1. Introduction

According to Holmberg's (2005, 2010) Null Generic Subject Generalization (NGSG), partial null subject languages (pNSLs) like Brazilian Portuguese have generic third-person null subjects, as in (1), while consistent null subject languages (cNSLs) like Spanish only allow referential third-person null subjects, as in (2) below.

(1) *Naquele quarto pro dorme bem.*
in-that bedroom one sleeps well
'In that bedroom one sleeps well.'

(2) *En ese cuarto pro duerme bien.*
in that bedroom s/he sleeps well
'In that bedroom s/he sleeps well.'

Holmberg attributes this difference to parametric variation localized on the T node as summarized in (3).

- (3) a. *Holmberg's (2005, 2010) Null Generic Subject Generalization (NGSG):*
 (i) cNSLs have a D-feature in T; pNSLs lack a D-feature in T.
 (ii) Since cNSLs have D-in-T, subject *pro* is always interpreted as referential.
 (iii) Since pNSLs lack D-in-T, subject *pro* is typically interpreted as generic.
 b. *Special Morphology Condition (SMC):*
 cNSLs allow generic null subjects only when licensed by “special overt morphology.”

Part (3b) of the NGSG, the Special Morphology Condition (SMC), accounts for when a generic null subject is licensed in cNSLs by morphological elements such as the Romance reflexive pronoun *se/si*. Nevertheless, the Spanish data in (4) below appear to pose a problem for the NGSG.

- (4) *Uno_i duerme bien, cuando uno_i/pro_i duerme en ese cuarto.*
 one sleeps well when one sleeps in that bedroom
 ‘One sleeps well when one sleeps in that bedroom.’

In the adjunct clause in (4), an overt pronoun is possible, but *pro* is preferred and is generically interpreted being coreferential with generic *uno* in the matrix clause.¹

In what follows I provide an account of the data in (4) above following Frascarelli (2007). Contrary to the NGSG, generic *pro* does occur in cNSLs and is licensed via topic, just like referential null subjects. Additionally, generic *uno* as a licenser of generic null subjects is not accounted for under the SMC and, since subject *pro* can be generic even in the presence of a D-feature in T, I propose that this aspect of the NGSG be dispensed with.

The format of this paper is as follows. In Section 2, I present Frascarelli's (2007) analysis of topic-licensing of referential *pro* in Italian and extend it to Spanish referential *pro*. Section 3 is my analysis, following Frascarelli, of the novel data with generic *pro* in Spanish. In Section 4, I show that while the SMC accounts for impersonal *se*, it cannot account for generic *uno*. Section 5 is a discussion of the implications of my analysis for the NGSG and Section 6 is the conclusion.

1. An alternative interpretation is possible: ‘One sleeps well when he sleeps in that bedroom.’

2. Topic identification of referential subject *pro*

2.1 Frascarelli (2007)

Frascarelli focuses on how the syntax-discourse interface assigns referential interpretation to third-person *pro*. She shows that a topic in the left periphery values *pro*'s features resulting in coreferential interpretation. Her cartographic approach to the left-periphery distinguishes three types of topics: (1) the Aboutness-shift Topic (A-topic), (2) the Contrastive Topic, and (3) the Familiar Topic.² The A-topic is the main concern here since it values *pro*'s features.

The following data show how the A-topic and *pro* relation works:³

- (5) *Il mio capo*_i ... *pro*_i è un exreporter...*pro*_i mi ha preso in simpatia ... *pro*_i è ... lunatico è capace che domani non gli sto simpatica e *pro*_i mi sbatte fuori... – poi c'è *M.F.*_k che è questo che appunto sta facendo tipo praticantato per poi andare a fare l'esame da giornalista...quindi *lui*_k c'ha quanto meno la garanzia che *pro*_k può rimanere lì finché *pro*_k non farà l'esame cioè *lui*_i poi gli deve scrivere le referenze...
- 'My boss_i ... *he*_i is a former reporter...and *he*_i likes me... *he*_i is ... moody, maybe tomorrow *he*_i does not like me and *he*_i fires me ... – then there is *M.F.*_k who is practicing for his exam as a journalist ... so... *he*_k has a guarantee that *he*_k will stay there till *he*_k has made the exam because *he*_i then must write a report...'

In (5), the first overt subject, *il mio capo*, is the topic that the following *pros* refer to and so they are coindexed. When a new A-topic is introduced, *M.F.*, it could identify *pro* but instead the speaker uses another overt pronoun, *lui*, and the subsequent *pros* are linked to it and to *M.F.* In the last sentence, another overt *lui* is introduced as a new A-topic referring again to *il mio capo*. In this case, *lui* is required. If a *pro* is merged instead, it will be identified with *M.F.* Thus, *pro* is always interpreted in relation to the closest A-topic. The interpretation of *pro* does not depend upon the agreement feature of the head licensing it but by agreement with the topic. Since all the subjects in question here have the same ϕ -features (masculine third-person singular) it is not ϕ -feature ambiguities that are the issue but rather topic ambiguity, encoded via an [Aboutness] feature.

Crucially, Frascarelli claims that topic-identification of *pro* is part of the core grammar (Agree) and it applies under specific structural conditions. Given that the topic is located in the left-periphery and *pro* is merged in Spec,*v*, this might

2. See also Reinhart (1981) and Lambrecht (1994) for the Aboutness Topic.

3. Adapted from Frascarelli (2007: 703).

pose an obstacle to the theory, since Agree occurs within phases per the Phase Impenetrability Condition (Chomsky 2001, 2004). Frascarelli deals with this by assuming *pro* is at the edge of the phase, Spec,*v*, where it is accessible to agree with the topic in ShiftP. After an A-topic is merged, its null counterpart is merged in later clauses, until a new A-topic is introduced. Consider the following extract from (5) above:

- (6) *Il mio capo_i è un exreporter ... pro_{-i} è lunatico...*
 the my boss is a former-reporter he is moody
 ‘My boss is a former-reporter... He is moody..’

In (6) the overt subject *il mio capo* in the first clause is coindexed with the *pro* in the second clause. In this case, its null copy is present in the second clause, valuing the φ -features on *pro*. Additionally, a [Person] feature and an [Aboutness] feature are valued on *pro* by the topic or its null copy.⁴ The structure of the null subject clause in (6) is represented in (7) below.⁵

- (7) [_{ShiftP} ~~*Il mio capo*~~_(φ , α , Pn) Shift' [_{TP} [_{vp} *pro*_(φ , α , Pn) [_{VP}]]]] ...

In (7), the line from the topic in Spec,Shift to *pro* in Spec,*v* represents Agree; i.e., *pro*'s [Person] and [Aboutness] features are valued by the topic. *Il mio capo* is struck through because this is a null copy of the overt topic first introduced in the prior clause, now base-generated in Spec,Shift. *Pro* is at the phase edge in Spec,*v*, allowing agreement to take place across phases. The same derivation takes place with each subsequent *pro*, as in (6), until a new A-topic is introduced.

Importantly, only preverbal subjects qualify as A-topics. Postverbal subjects are associated with Focus (Frascarelli 2007: 726), which can be coreferential with the null A-topic. Thus, postverbal *by*-phrases should not count as topics. This is supported by the following Italian data from Samek-Lodovici (1996):

- (8) a. *Questa mattina, la mostra è stata visitata da Gianni_i. Più tardi egli_j/*
 this morning the exhibit is been visited by Gianni more late he
*lui_j/*pro_i ha visitato l'università.*
 he he has visited the-university
 ‘This morning, the exhibit was visited by Gianni. Later, he visited the university.’

4. “Person” in the logophoric sense.

5. Aboutness = α ; Person = Pn.

- b. *Questa mattina, Gianni, ha visitato la mostra. Più tardi pro_i ha visitato l'università.*
 this morning Gianni has visited the exhibit more late he has visited the-university
 'This morning, Gianni visited the exhibit. Later, he visited the university.'

In (8a), the preverbal subject of the first sentence is *la mostra* and not *Gianni*, which occurs in a *by*-phrase. As a result, a *pro* referring to *Gianni* in the following sentence is ungrammatical. In (8b), however, *pro* is acceptable in the second sentence because *Gianni* is now a preverbal subject that serves as a topic.

2.2 Extension of Frascarelli (2007) to Spanish referential *pro*

The NGSg predicts that null subjects are licensed identically in cNSLs. Thus, the analysis given in Section 2.1. for Italian *pro* should apply to Spanish. This is supported by the following data, a Spanish adaptation of (8) above.

- (9) a. *Esta mañana, la exhibición fue visitada por Juan. Luego, él/*pro_i fue a la universidad.*
 this morning the exhibit was visited by John later he went to the university
 'This morning, the library was visited by John. Later, he went to the university.'
- b. *Esta mañana, Juan_i visitó la exhibición. Luego, pro_i fue a la universidad.*
 this morning John visited the exhibit later he went to the university
 'This morning, John visited the library. Later, he went to the university.'

In (9a), the preverbal subject is *la exhibición*; it is also the A-topic identifying *pro*. In the second sentence of (9a), *pro* is unacceptable, but *él* is acceptable. The null subject is not acceptable because it can only be identified with *la exhibición*. Thus, (9a) shows the same pattern as Italian (8a) above. In (9b), *pro* is acceptable in the second sentence because *Juan*, the preverbal subject of the preceding sentence, is also the topic. Consequently, its null copy is present in later clauses and can identify *pro*. (9b) shows the same patterns as Italian (8b) above, with the same type of agreement taking place as below.

- (10) [ShiftP Juan_i_[φ,α, Pn] Shift' [TP [_{vp} PRO_[φ,α, Pn] [_{vp} fue a la universidad]]]]]

In (10), the null topic *Juan* is in Spec,Shift while *pro* is merged in Spec,*v*, at the edge, allowing Agree. The A-topic values the relevant features on *pro*, thereby ensuring

the same interpretation. In conclusion, Spanish referential *pro* is identified by A-topic, just like Italian referential *pro* following Frascarelli's (2007) analysis.

3. Topic identification of generic *pro*

3.1 Topics can be indefinite

Recall the data in (4) repeated as (11) below:

- (11) *Uno_i duerme bien, cuando uno_i/pro_i duerme en ese cuarto.*
 one sleeps well when one sleeps in that bedroom
 'One_i sleeps well when one_i sleeps in that bedroom.'

In (11), the subject in the second clause may be null and, being coreferential with *uno*, it is also generic, a situation predicted to not be possible given Holmberg's NGSG. I propose that in order for *pro* to receive a generic interpretation, as with referential *pro*, it enters into an agreement relationship with the local A-topic which, in this case, is generic *uno*. As is shown in (12) below, in the matrix clause, *uno* merges first in Spec,*v* but moves to Spec,T for EPP and later to Spec,Shift, since it is the A-topic.

- (12) [_{ShiftP} *uno* [_{TP} <*uno*> [_{VP} <*uno*> [_{VP} *duerme bien*]]]]

In the adjunct clause, the null subject is at the phase edge where it can have its [Aboutness] and [Person] features valued by the topic, across its own phase. The structure of the adjunct clause in (11) is represented below.

- (13) [_{CP} *Cuando* _{ShiftP} ~~*uno*~~_i [_{TP} *pro*_i [_{VP} <*pro*>_i [_{VP} *duerme en ese cuarto*]]]]
└──────────┘
 AGREE

In (13), *pro* merges in Spec,*v*; this is where Agree with the null topic in Spec,Shift takes place. Afterward, *pro* is free to move to Spec,T for EPP.

Holmberg et al.'s (2009: 70) analysis of null referential *pro* in cNSLs stipulates that A-topics are always "definite." They make this claim to account for the valuing of a D-feature in T. However, this may not be necessary. Indeed, while Holmberg et al. (2009) base their study on Frascarelli (2007), they seem to disagree with her with respect to the definiteness of topics. Consider the following data from Italian, which Frascarelli provides as an example of topic-identification of *pro*.⁶

6. (14) is an excerpt from Frascarelli's (2007: 705) example (14).

- (14) *Mi sono fermata... a un benzinaio e gli ho detto, 'Scusi me am stopped at a gas-station-attendant and him have said pardon ma se uno_p poveraccio, sbaglia qui, come pro_i fa a tornare indietro?'*
 but if one poor-guy goes-wrong here how one makes to return back
 'I stopped... at a gas station attendant and said to him, 'Excuse me, but if one, poor guy, goes the wrong way here, how does he get back?'

Frascarelli (2007: 706) states that the *uno* in (14) is a DP “produced as an Aboutness-shift Topic” that “can provide a value for the null subject (*pro*) in the following sentence.” Thus, for Frascarelli, indefinite *uno* is a topic. Given that Holmberg et al.'s (2009) analysis is based on Frascarelli (2007), it is surprising that this assumption is not shared.⁷

Interestingly, Frascarelli's Example (14) above contains the same generic pronoun under consideration here, *uno*, which is homophonous in both languages. In Spanish, just like Italian, *uno* can be a topic, as from the following data show:

- (15) *Cuando uno_i es rico, pro_i viaja frecuentemente. pro_i Viaja a Francia o a Italia y pro_i bebe vino caro.*
 when one is rich one travels frequently one travels to France or to Italy and one drinks wine expensive
 'When one is rich, he travels frequently. One travels to France or to Italy and drinks expensive wine.'

In (15), *uno* as a topic is the antecedent of every occurrence of *pro*, even in an independent sentence. In this way, it patterns with definite topics as seen above in (5). Consequently, we may conclude that *uno* can be a topic in at least two cNSLS, Spanish and Italian. As such, it can identify *pro*'s content in subsequent clauses.

3.2 An ordering constraint

Further evidence that generic *uno* functions as a topic comes from different orders between matrix and adjunct clauses containing *uno* and coreferential *pro*. With respect to the data in (11) above, there are four ordering possibilities. These different configurations are represented in the data below.

7. Other types of indefinites also license a *pro* in later clauses, as shown in the following Spanish data:

- (i) *Algunas personas_i prefieren un asiento de pasillo cuando pro_i viajan en avión.*
 some people prefer a seat of aisle when they travel on plane
 'Some people prefer an aisle seat when they travel by plane.'
- (ii) *Un buen estudiante_i no se queja cuando pro_i saca malas notas.*
 a good student not Pron_{se} complains when he gets bad grades
 'A good student does not complain when he gets bad grades.'

- (16) *Uno_i duerme bien, cuando pro_i duerme en ese cuarto.*
 one sleeps well when one sleeps in that bedroom
 ‘One_i sleeps well when one_i sleeps in that bedroom.’
- (17) **Pro_i duerme bien, cuando uno_i duerme en ese cuarto.*
 one sleeps well when one sleeps in that room
- (18) *Cuando uno_i duerme en ese cuarto, pro_i duerme bien.*
 when one sleeps in that room one sleeps well
 ‘When one sleeps in that room, one sleeps well.’
- (19) **Cuando pro_i duerme en ese cuarto, uno_i duerme bien.*⁸
 when one sleeps in that room one sleeps well

These data show that *uno* must precede *pro* in the discourse. In (16), matrix *uno* precedes the adjunct *pro* while in (18) sentence-initial adjunct *uno* precedes matrix *pro*. Clause type is irrelevant since *uno* occurs as both matrix and adjunct subject and the same interpretation obtains. What is important is discourse order. In (17), matrix *pro* precedes adjunct *uno* and the sentence is ungrammatical. In (19), adjunct *pro* precedes matrix *uno* and the same ungrammaticality occurs.

This pattern is predicted by the analysis given in Section 3.1. above, where *uno* is taken to be a topic that values the relevant features on *pro*. A null copy of *uno* is present in the following clauses, carrying out the same valuation until a new A-topic is introduced. In (16) and (18), a null copy is present in the second clause because *uno* has already been established as the topic in the first clause. This is represented below with (20) for (16) and (21) for (18).

- (20) a. Matrix

$$[\text{ShiftP } \text{uno}_i [\text{TP } \langle \text{uno} \rangle_i [\text{VP } \langle \text{uno} \rangle_i \text{ duerme bien }]]]$$
- b. Adjunct

$$[\text{CP } \text{cuando } [\text{ShiftP } \text{uno}_i [\text{TP } [\text{pro}_i [\text{VP } \langle \text{pro} \rangle_i \text{ duerme ... }]]]]]$$
- AGREE
- (21) a. Adjunct


$$[\text{CP } \text{cuando } [\text{ShiftP } \text{uno}_i [\text{TP } \langle \text{uno} \rangle_i [\text{VP } \langle \text{uno} \rangle_i \text{ duerme ... }]]]]$$

8. A reviewer points out that the referential version of (19) is acceptable:

- (i) *Cuando pro_i duerme en ese cuarto, el niño_i duerme bien.*

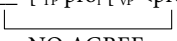
Mara Frascarelli has accounted for this via personal communication as a result of *el niño* in the second clause being a G-topic located in the Ground Phrase, which accommodates the presence of a null topic. The generic version in (19) is unacceptable because generic *uno* cannot be a G-topic. Generic *uno* can be an A-topic, as I have argued above in Section 3.2; see also Reinhart (1981).

- b. Matrix

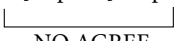
$$[\text{ShiftP } \underline{\text{uno}}_i [\text{TP } [\text{pro}_i [\text{VP } \langle \text{pro} \rangle_i \text{ duerme bien }]]]]]$$


As is shown in (20b) and (21b), a null copy of *uno* is in Spec,Shift agreeing with *pro* and thus the sentence is grammatical. When no null copy is present, ungrammaticality results, as shown below where (22) represents (17) and (23) is (19).

- (22) a. Matrix

$$* [\text{ShiftP } \underline{\quad} [\text{TP } \text{pro}_i [\text{VP } \langle \text{pro} \rangle_i \text{ duerme bien }]]]]$$

- b. Adjunct

$$[\text{CP } \text{cuando } [\text{ShiftP } \text{uno}_i [\text{TP } \langle \text{uno} \rangle_i [\text{VP } \langle \text{uno} \rangle_i \text{ duerme ... }]]]]]$$
- (23) a. Adjunct

$$* [\text{CP } \text{cuando } [\text{ShiftP } \underline{\quad} [\text{TP } \text{pro}_i [\text{VP } \langle \text{pro} \rangle_i \text{ duerme ... }]]]]]]$$

- b. Matrix

$$[\text{ShiftP } \text{uno}_i [\text{TP } \langle \text{uno} \rangle_i [\langle \text{uno} \rangle_i \text{ duerme bien }]]]]]$$

In (22a) and (23a) there is no topic in Spec,Shift identifying *pro* as represented by the underscore. *Uno* is an eligible topic but it is introduced in the second clause, where it is not in a position to value *pro*'s [Aboutness] and [Person] features in the first clause. Thus, features go unvalued and the derivation crashes.

4. The special morphology condition: Impersonal *se* and generic *uno*

Recall that the SMC allows for null generic subjects in cNSLs when licensed by special morphology such as Spanish impersonal *se* (Imp_{se}), as in (24).

- (24) *En ese cuarto, se pro duerme bien.*
 in that room Imp_{se} one sleeps well
 'In that room, one sleeps well.'

Authors like Otero (1986), Mendikoetxea (2008), MacDonald (2017) have analyzed Spanish Imp_{se} as in (24), as the head, *v* or Voice, that licenses generic *pro*. Thus, Imp_{se} qualifies as part of the SMC. In this section, I consider whether generic *uno* also falls under this condition.

4.1 Coreferentiality diagnostics

Patterns of cross-clausal coreferentiality can be used to diagnose whether generic *uno* and *Imp_{se}-pro* are similar elements. If they pattern together, this may be taken as positive evidence. Consider the following data.

- (25) *Cuando uno_i trabaja duro, se_i pro_i gana mucho dinero.*
 when one works hard *Imp_{se}* one earns much money
 ‘When one works hard, one earns a lot of money.’

In (25), *uno* and *Imp_{se}-pro* appear to be coreferential across clauses.⁹ Additionally, the order can be reversed as shown below in (26).

- (26) *Cuando se pro_i trabaja duro, uno_i gana mucho dinero.*

Imp_{se}-pro can be coreferential with *Imp_{se}-pro*, as in (27).

- (27) *Cuando se pro_i trabaja duro, se pro_i gana mucho dinero.*

However, *Imp_{se}-pro* and bare third-singular *pro* cannot be coreferential, as seen in (28) and (29).

- (28) **Cuando se pro_i trabaja duro, pro_i gana mucho dinero.*

- (29) **Cuando pro_i trabaja duro, se pro_i gana mucho dinero.*

Now compare (28) with (30), repeated from above.

- (30) *Cuando uno_i trabaja duro, pro_i gana mucho dinero.*

These data show that *Imp_{se}* and generic *uno* display some similarities with respect to cross-clausal coreferentiality, but they differ in a critical way: *Imp_{se}* does not license bare *pro* in later clauses, while generic *uno* does.

4.2 Uno can be an A-topic; *Imp_{se}* cannot be an A-topic

The different patterns seen in above can be attributed to the inability of *Imp_{se}* to be a topic. Recall the data in (31), repeated from above:

9. A reviewer points out that while (25) is acceptable, other instances of *uno* and *se* are not:

- (i) *Cuando uno trabaja duro, (*se) pro duerme poco.*
 ‘When one works hard, one sleeps little.’
- (ii) *Cuando uno trabaja duro, (*se) pro es recompensado.*
 ‘When one works hard, one is rewarded.’

The reviewer takes this to mean that *uno* cannot identify *Imp_{se}-pro* via topic identification, which is consistent with my argument. Generic *uno* identifies a bare generic *pro* in later clauses while *Imp_{se}* can only license generic *pro* locally.

- (31) *Cuando uno_i es rico, pro_i viaja frecuentemente. pro_i Viaja a Francia o a Italia y pro_i bebe vino caro.*
 when one is rich one travels frequently one travels to France or to Italy and one drinks wine expensive
 ‘When one is rich, one travels frequently. One travels to France or to Italy and drinks expensive wine.’

In (31) generic *uno* identifies generic *pro* in subsequent clauses as an A-topic, functioning even across full stops. Compare (31) with (32) below, where generic *uno* is replaced with *Imp_{se}*.

- (32) **Cuando se_i pro_i es rico, pro_i viaja frecuentemente. pro_i Viaja a Francia o a Italia y pro_i bebe vino caro.*
 when *Imp_{se}* one is rich one travels frequently one travels to France or to Italy and one drinks wine expensive
 ‘When one is rich, he travels frequently. One travels to France or to Italy and drinks expensive wine.’

In (32), *pro* is ungrammatical because it is not identified in later clauses by *Imp_{se}*, unlike what was seen in (31) with generic *uno*. Now compare (32) with (33) below.

- (33) *Cuando se_i pro_i es rico, se pro_i viaja frecuentemente. Se pro_i viaja a Francia o a Italia y se pro_i bebe vino caro.*
 when *Imp_{se}* one is rich *Imp_{se}* one travels frequently *Imp_{se}* one travels to France or to Italy and *Imp_{se}* one drinks wine expensive
 ‘When one is rich, he travels frequently. One travels to France or to Italy and drinks expensive wine.’

(33) shows that (32) is ameliorated by inserting an *Imp_{se}* local to every occurrence of *pro*. Thus, *pro* can be licensed by *Imp_{se}*, but only locally. *Uno* can also license *pro*, but it is not restricted by locality. This difference is due to the inability of *Imp_{se}* to be an A-topic. *Imp_{se}*, as the spell out of little *v* or Voice, licenses *pro* (Mendikoetxea 2008, MacDonald 2017) but since it is a functional head it cannot serve as an A-topic. In *Imp_{se}*-constructions, a generic operator binds *pro* and T, leading to a generic interpretation. Generic *uno*, on the other hand, is a full pronominal and not a functional head. Consequently, it is a suitable candidate for an A-topic, which is how it licenses *pro* in subsequent clauses.¹⁰

10. A reviewer points out that the following is acceptable:

- (i) *La gente disfruta mucho en los viajes. Se baila, se canta y se conoce a nuevos amigos.*
 ‘People enjoy themselves a lot on trips. They dance, sing, and meet new friends.’

5. Implications for the null generic subject generalization

Holmberg (2005) observes that cNSLs do not allow null generic subjects while pNSLs do. I have shown in the preceding sections that in fact, contra Holmberg's claim, cNSLs do have generic null subjects under conditions unrelated to a D-feature in T; i.e., the presence of a generic topic. This raises the question as to whether there is any reason to maintain the D-in-T aspect of the NGSG. Holmberg and his coauthors build their analyses of the various interpretations of *pro* around the presence or absence of D-in-T. However, in Sections 3 and 4 I demonstrated that the D-feature is not needed in order to account for either referential or generic *pro* in cNSLs. Thus, I postulate that the same interpretations of *pro* can be derived in pNSLs without using the D-feature in T.

5.1 Referential *pro* in pNSLs

Referential subject *pro* in pNSLs is also licensed via topic-identification. In pNSLs such as BP, a referential null subject is only allowed when controlled by a higher argument, as in (34).

- (34) *O João_i disse que ele_i/pro_i tinha comprado uma casa.*¹¹
 the João said that he had bought a house
 'João said that he had bought a house.'

In (34), the embedded *pro* is controlled by the matrix subject *O João*. Holmberg et al. (2009) and Holmberg and Sheehan (2010) stipulate in their analysis that this *pro* has an unvalued D-feature that is valued by the D-feature on the controller argument. This stipulation is necessary for them because, following the NGSG, BP has no D-feature in T, and having a D-feature somewhere is how a null subject is interpreted as referential. However, the data in (34) are very similar to the patterns seen in Section 2 for referential *pro* in cNSLs; i.e., a previous overt argument identifies the referent of *pro*. Hence, I propose that the same analysis given for referential *pro* in cNSLs in Section 2 can be applied to referential *pro* in pNSLs; i.e., topic identification. This analysis is represented in (35) below.

This is not surprising given my account since the only *pros* involved are in the second sentence which all have an Imp_{se} present locally to license them. Nevertheless, this does suggest that *la gente* can also be a topic licenser of bare generic *pro* as below:

- (ii) *La gente disfruta mucho en los viajes. Baila, canta, y conoce a nuevos amigos.*
 'People enjoy themselves a lot on trips. They dance, sing, and meet new friends.'

11. This example taken from Holmberg et al. (2009: 65)

- (35) a. Matrix
 [ShiftP *O João*_i [TP <*O João*>_i [vP <*O João*>_i *disse*]]]
 b. Embedded
 [CP *que* [ShiftP *O João*_i TP [vP *pro*_i *tinha comprado* . . .]]]

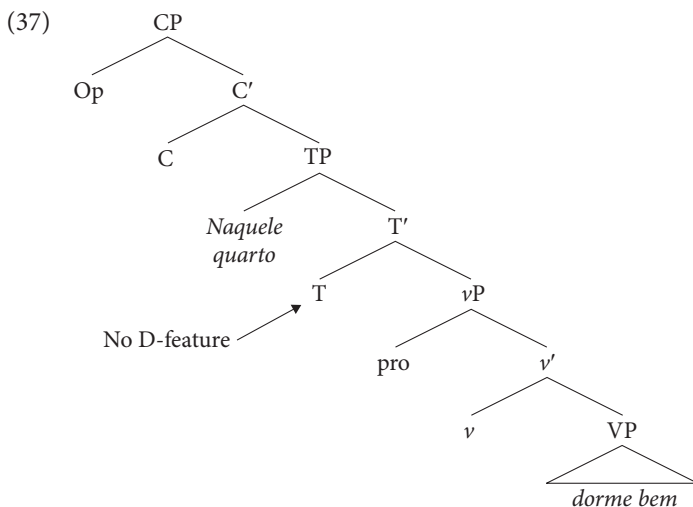
In (35) the preverbal subject of the matrix clause, *O João*, is the A-topic. Its null copy is base-generated in Spec,Shift in the embedded clause, identifying *pro* through agreement.

5.2 Generic *pro* in pNSLs

Holmberg's (2005, 2010) analysis of generic *pro*, following Moltmann (2006), can be adopted with little modification, since it rests on the absence of a D-feature, which is something I have already proposed above. In pNSLs, third-person null subjects are generic because there is no D-feature in T. Recall the BP data in (36) repeated from above.

- (36) *Naquele quarto pro dorme bem.*
 in-that bedroom one sleeps well
 'In that bedroom one sleeps well.'

A Holmberg-type analysis of the data in (36) is represented in (37) below.



In (37), a generic operator in Spec,C binds T and *pro*, resulting in generic interpretation. Since this is a pNSL, there is no D-feature present to make *pro* a referential. The locative topic, *naquele quarto*, is required in Spec,C in order to check the EPP.

In this way, a topic still has a role to play in the derivation, but there is no Agree between it and *pro*.

5.3 Revising the NGSG

Holmberg's original observation that cNSLs differ with respect to the distribution of null generic subjects is accurate. However, on my account, these differences are not related to a D-feature. Thus, I propose that the NGSG be revised as below in (38).

(38) *Null Generic Subject Generalization (REVISED)*:

- a. cNSLs and pNSLs both allow referential and generic *pro*; they differ with respect to licensing conditions.
- b. In cNSLs, referential and generic *pro* are licensed via agreement with a null referential or generic A-topic, respectively.
- c. In pNSLs, referential *pro* is licensed via agreement with a null referential topic. Generic *pro* licensed by a generic operator in Spec,C accompanied by a locative phrase in Spec,T for EPP.

6. Conclusion

I have shown that, contrary to Holmberg's (2005, 2010) claim, consistent null subject languages (cNSLs) such as Spanish and Italian do allow null generic subjects. In these languages, null generic *pro* is licensed via topic-identification. A null copy of a generic topic enters into Agree with *pro*, forming a chain that results in generic interpretation. Unlike impersonal *se*, generic *uno* is not accounted for under Holmberg's Special Morphology Condition, since the former only licenses generic *pro* locally while the latter can license generic *pro* throughout the discourse. Finally, since generic null subjects are licensed in cNSLs in a way unrelated to the presence or absence of a D-feature in T, the revised version of the Null Generic Subject Generalization given above in (38) captures the different crosslinguistic distribution and licensing conditions on null subjects without making reference to this feature. These findings contribute to our understanding of the licensing of null subjects crosslinguistically and lend further support to the topic-identification analysis proposed by Frascarelli (2007).

Acknowledgements

I thank the audience at LSRL 2016 and the reviewers for helpful comments. I am grateful to Mara Frascarelli, Jonathan MacDonald, Vieri Samek-Lodovici, Amanda Barón, Renato Lacerda, Jeriel Melgares, Cristina Mostacero, and Begoña Aretxabaleta. All errors are my own.

References

- Chomsky, Noam. 2001. "Derivation by Phase." In *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 2004. "Beyond Explanatory Adequacy." In *Structures and Beyond. The Cartography of Syntactic Structures, Volume 3*, ed. by Adriana Belletti, 104–131. New York: Oxford University Press.
- Frascarelli, Mara. 2007. "Subjects, Topics and the Interpretation of Referential Pro: An Interface Approach to the Linking of (Null) Pronouns." *Natural Language and Linguistic Theory* 25: 691–734. <https://doi.org/10.1007/s11049-007-9025-x>
- Holmberg, Anders. 2005. "Is There a Little Pro? Evidence from Finnish." *Linguistic Inquiry* 36: 533–564. <https://doi.org/10.1162/002438905774464322>
- Holmberg, Anders. 2010. "The Null Generic Subject in Finnish: A Case of Incorporation in T?" In *Parametric Variation. Null Subjects in Minimalist Theory*, ed. by Theresa Biberauer, Anders Holmberg, Ian Roberts, and Michelle Sheehan, 200–230. Cambridge: Cambridge University Press.
- Holmberg, Anders, Aarti Nayudu, and Michelle Sheehan. 2009. "Three Partial Null-Subject Languages: A Comparison of Brazilian Portuguese, Finnish and Marathi." *Studia Linguistica* 63: 59–97. <https://doi.org/10.1111/j.1467-9582.2008.01154.x>
- Holmberg, Anders, and Michelle Sheehan. 2010. "Control into Finite Clauses in Partial Null-Subject Languages." In *Parametric Variation. Null Subjects in Minimalist Theory*, ed. by Theresa Biberauer, Anders Holmberg, Ian Roberts, and Michelle Sheehan, 125–152. Cambridge: Cambridge University Press.
- Lambrecht, 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>
- MacDonald, Jonathan E. 2017. "An Implicit Projected Argument in Spanish Impersonal and Passive Se Constructions." *Syntax* 20 (4): 353–383.
- Mendikoetxea, Amaya. 2008. "Clitic Impersonal Constructions in Romance: Syntactic Features and Semantic Interpretation." *Transactions of the Philological Society* 106: 290–336. <https://doi.org/10.1111/j.1467-968X.2008.00210.x>
- Moltmann, Friederike. 2006. "Generic One, Arbitrary PRO, and the First Person." *Natural Language Semantics* 14: 257–281. <https://doi.org/10.1007/s11050-006-9002-7>
- Otero, Carlos P. 1986. "Arbitrary Subjects in Finite Clauses." In *Generative Studies in Spanish Syntax*, ed. by Iyonne Bordelouis, Heles Contreras, and Karen Zagona, 81–109. Dordrecht: Foris.
- Reinhart, Tanya. 1981. "Pragmatics and Linguistics: An Analysis of Sentence Topics." *Philosophica* 27: 53–94.
- Samek-Lodovici, Vieri. 1996. *Constraints on Subjects: An Optimality Theoretic Analysis*. Ph.D. dissertation, Rutgers University.

Bridging and dislocation in Catalan

Laia Mayol and Xavier Villalba

Universitat Pompeu Fabra / Universitat Autònoma de Barcelona

The exact anaphoric relation between the referent of dislocates and their previous discourse referents has been a matter of much discussion. In this paper, we reexamine this relation under the light of Clark's notion of bridging (Clark 1977), and connect different anaphoric bridging types with left (LD) and right dislocations (RD) in Catalan. We report a judgment task experiment showing that while speakers clearly perceive LD and RD as different for most kinds of bridging, they rate RD better across the board, suggesting that speakers can use RD for a wider range of uses than previously assumed. Yet, albeit the preference for RD in Catalan, LD seems to be inherently linked to contrastive readings.

Keywords: bridging, left dislocation, right dislocation, anaphora, contrast, Catalan

1. Introduction

Clark (1977) describes bridging as a strategy for accommodating the given-new structure in cases where the given content is given for the speaker, but it is not for the hearer:

- (1) *John saw someone leaving the party early. It was Mary who left.*
- (2) *In the group there was one person missing. It was Mary who left.*

In the first utterance in (1), the speaker explicitly mentions that someone left, so the following cleft can be used to complete the information with Mary, while presupposing that someone left. Compare with the sequence in (2). Here the first utterance provides less information: "X was missing". However, the speaker can felicitously utter the cleft in the second utterance, just as in (1), taking as presupposed "someone left", even though this information is not provided in the previous discourse: we only know that "someone was missing". In this case, the hearer builds a plausible inference (i.e. a bridge) from "someone was missing" to "Mary left".

Bridging is pervasive in language and has been studied thoroughly (see Matsui 1993, Asher and Lascarides 1998, Irmer 2011, a.o.), but it figures prominently in the realm of dislocation structures, under different labels: partial ordered set (*poset*) relations (Ward and Prince 1991), anaphoricity (López 2009) or monotonicity (Bott 2007). In this paper, we test the connection between bridging phenomena and clitic left- (LD) and right-dislocation (RD) in Catalan by means of an experimental study. The paper is structured as follows: Section 2 presents some background on both bridging and dislocation constructions. Section 3 outlines the materials and methods of the experimental design. Section 4 presents the results and Section 5 discusses them. Finally, Section 6 summarizes the main conclusions.

2. Background

2.1 Bridging

As mentioned, bridging is considered by Clark as a type of inference necessary to accommodate particular given-new structures. Clark considers that this inference is a species of Gricean conversational implicature, but one involving not just linguistic knowledge, but world knowledge and accommodation as well. If we return to our example in (2), it is clear that the speaker must make an inference from known information available in discourse, but also from world knowledge, just as happens with conversational implicatures (Grice 1975, 1981, 1989). Finally, bridging needs to accommodate the inferred meaning, for it is taken as granted even though it is not part of the common ground (on accommodation, see Lewis 1979, Stalnaker 2002, 2014, Simons 2003). In this sense, bridging is akin to presupposition.

Clark discusses different types of bridging relationships, the most important for our purposes being the following:¹

1. Direct reference: the given information makes direct reference to something just mentioned.
 - a. Identity: “The house was large. The size surprised me.”
 - b. Pronominalization: “I met a man yesterday. He told me a story.”
 - c. Epithets: “I met a man yesterday. The bastard stole all my money.”
 - d. Set membership: “I met two doctors yesterday. The tall one told me a story.”

1. There is a fourth type of bridging relationship discussed in Clark (1977), in which the given information provides Reasons, Causes, Consequences or Concurrences of previously mentioned states or events. We have not included this category in our study for it usually involves VP, rather than nominal, anaphora.

2. Indirect reference by association: the given information does not refer to something explicitly mentioned, but it is closely associated to something previously mentioned.
 - a. Necessary parts: “I looked into the room. *The ceiling* was very high.”
 - b. Probable parts: “I walked into the room. *The windows* looked out to the bay.”
 - c. Inducible parts: “I walked into the room. *The chandeliers* sparkled brightly.”
3. Indirect reference by characterization: the given information plays a role in an event previously mentioned in the discourse.
 - a. Necessary roles: “John was murdered yesterday. *The murderer* got away.”
 - b. Optional roles: “John died yesterday. *The murderer* got away.”

2.2 Left and right dislocation

So far, we have covered the basics of bridging phenomena. Now, we will consider the relation between bridging and dislocation constructions.

Since Ward (1988) (see also Ward and Prince 1991, and Birner and Ward 1998), it has been commonly assumed that topicalized phrases are licensed under a *partial ordered set* (*poset*) relation with a previous discourse referent:²

Discourse Condition on Preposing in Topicalization

The entity represented by the preposed constituent must be related, via a salient partially ordered set relation, to one or more entities already evoked in the discourse model. (Ward and Prince 1991: 173)

Vallduví (1992) extends the poset analysis to left-dislocations (LD, henceforth), or to be more precise, to *links* in general, which are pointers to a certain address in the hearer knowledge store:

In other words, hearers cannot jump from one address to the other unless those two addresses are related via a poset relation. The ban on the preposing of constituents that denote addresses that fail to be in a poset relation with some already-evoked address is, then, a reflection of the fact that the address the hearer is instructed to go to is not accessible from the address s/he is at the time of utterance. Only addresses that are in a poset relation with the current address are accessible. (Vallduví 1992: 91)

However, successful as it may seem for LD (see Villalba 2009: Chapter 2, for a survey of the whole range of cases included under the poset relation), the poset

2. A partially ordered set is a set defined by a relation that is either reflexive, antisymmetric and transitive or irreflexive, asymmetric, and transitive. See Wall (1972: 141).

account has to face a problematic case: inferable relations (see also Hendriks 1996, Hendriks and Dekker 1995). Observe the following dialogues, from Villalba (2009: Chapter 2), where a bridging relation is established by means of LD:

- (3) a. *Els obrers començaran a treballar a l'edifici a les vuit.*
 'The workers will begin to work in the building at eight o'clock.'
 b. *Doncs als veïns no els farà gens de gràcia que els despertin tant d'hora.*
 'Well, the neighbors won't be amused with being awakened so early.'
- (4) a. *M'agrada molt aquesta casa.*
 'I like this house very much.'
 b. *A l'arquitecte, en canvi, no el va deixar satisfet el resultat.*
 'The architect, however, wasn't satisfied with the result.'

Even though LD is perfect in these contexts, no clear poset relation can be established between *edifici* 'building' and *veïns* 'neighbors' or *casa* 'house' and *arquitecte* 'architect', even though the relation is evident to everyone. Ward's *Discourse Condition on Preposing in Topicalization* does not formalize this possibility. Indeed, these are clear instances of bridging: in the former case, building and neighbors are in an optional role bridging relation, whereas in the later, the relation seems rather one of necessary role.

While these cases are fairly well described for left-detachments (Hendriks and Dekker 1995, Bott 2007, Brunetti 2009), right-dislocations (RD, henceforth) have received much less attention. To our knowledge, it was Ziv (1994) who originally remarked that right-dislocates involved bridging anaphora in cases like the following (see also Grosz and Ziv 1998):

- (5) a. *I saw Modern Times again yesterday.*
 b. *He's amazing, (this) Charlie Chaplin.*

In this case, the right dislocate is inferred from the previous mention of the movie, and hence it is a clear case of bridging.

These cases are a problem for an approach like the one defended in Bott (2007), namely that RDs (tails) must be upward-monotonic anaphors, whereas LDs (links) must be downward anaphors or non-monotonically anaphoric. He offers the following examples for Catalan, which we have adapted for ease of exposition:

- (6) a. *Which relationship did Bach have to string instruments?*
 b. *[La viola]_{LINK}, segur que li va agradar.*
 'The viola, he surely liked.'
 c. *#Segur que li va agradar, [la viola]_{TAIL}.*
 'He surely LIKED, the viola.'

- (7) a. *Which relationship did Bach have to the viola?*
 b. *Segur que li van agradar, [els instruments de corda]_{TAIL}*
 ‘He surely LIKED, string instruments.’
 c. *#[Els instruments de corda]_{LINK} segur que li van agradar.*
 ‘String instruments, he surely liked.’

In (6), one can appreciate that LD (*the viola*) is felicitous when it is more specific than its antecedent (*string instruments*): we move from a general antecedent to a more specific dislocate – downward anaphora, then. In (7), we reverse the situation with RD (*string instruments*), which is more general than its antecedent (*the viola*): we move from a specific antecedent to a more general dislocate – upward anaphora, then.

Yet, Bott acknowledges the existence of bridging cases not easily accounted under such a proposal (i.e. (3)–(5)), which he tries to reduce them to a subspecies of part-of relations, hence of downward-anaphora. Obviously, this makes the prediction that bridging anaphora would be fine for links, but not for tails, which is plainly wrong.

The view represented by Bott, namely that the conditions licensing RD are stricter than those of LD, is already present in Villalba (2000: Chapter 3), who shows that RDs are only possible in a subset of the poset relations, most notably identity relations. Even though he acknowledges the bridging cases by Ziv (1994) just mentioned, his main conclusion is that the closer the relation between the antecedent and the dislocate (identity being the closest one), the better the RD (see Brunetti 2009 for similar conclusions for Italian).

To sum up, even though the literature on dislocation has identified several anaphoric relations to hold between dislocates and their discourse antecedents, which could fall under the label of bridging, we still don't have a clear view of the connection between these two phenomena. Our goal is to determine which information functions of LDs and RDs are better suited for encoding different bridging phenomena, and particularly whether the closeness of the anaphoric linkage between the antecedent and the dislocate is related to the choice between LD and RD.

3. Testing bridging and dislocation

3.1 Experimental design

We designed a judgment task with two dislocation types (LDs and RDs) and seven bridging types: HYPONYM (*fruit-watermelon*), SET MEMBERSHIP (*Italian cities-Florence*), EPITHET (*John-that idiot*), NECESSARY PARTS (*train-wagon*), INDUCIBLE PARTS (*kitchen-coffee maker*), OPTIONAL ROLE (*death-murder*), and NECESSARY

ROLE (*murder-murderer*). Each item contained a sequence of two sentences: the second sentence was presented either with a left or a right-dislocation and the dislocate phrase included given information that needed to be bridged to an antecedent in the first sentence. Example (8) shows an example for the INDUCIBLE PARTS type (*flat-heater*) in the RD condition. In the LD condition (see (9)), the two sentences were identical except for the fact that the phrase “els radiadors” would be presented as a LD.

- (8) a. *La meva cosina s'ha canviat de pis i ahir va fer una festa.*
 ‘My cousin moved to another flat, and yesterday she organized a party.’
 b. *Els ha pintat de color taronja, els radiadors.*
 them has painted of color orange, the.pl heaters
 ‘She has painted the heaters orange.’
- (9) a. *La meva cosina s'ha canviat de pis i ahir va fer una festa.*
 ‘My cousin moved to another flat, and yesterday she organized a party.’
 b. *Els radiadors, els ha pintat de color taronja.*
 the.PL heaters them has painted of color orange
 ‘She has painted the heaters orange.’

When building the items, we were very careful to place the antecedent in focus position (the rightmost position in the main clause), and we opted for object dislocates to avoid unwanted readings, for Catalan object dislocates must be resumed by a clitic (unlike subjects), and objects are not clitic-doubled (unlike datives).

Note that three of the types (namely, HYPONYM, SET MEMBERSHIP and EPI-THET) are clear instances of poset relations, while the rest are not. Among the non-poset relations, they all represent bridging by indirect reference: two of the types represent a closer bridging relationship (NECESSARY PARTS and NECESSARY ROLES), while the other two represent more distant bridging relationships (INDUCIBLE PARTS and OPTIONAL ROLE).

Two counterbalanced randomized lists were prepared with 42 target items (= (3 sentences with LD + 3 sentences with RD) × 7 bridging types) and 40 fillers. Since the experiment was posted in the web (Survey Monkey), it was preceded by a language proficiency questionnaire and a brief instruction section. 168 native Catalan speakers completed the experiment, in which they had to rate the acceptability of every item in a 10-point Likert scale, as shown in Figure 1. We choose a 10-point scale for two main reasons: (i) we wanted to avoid having a scale with a midpoint, and (ii) a 10-point scale is intuitive for our participants (mostly university students), since essentially the same system is used for grades. Participants evaluated one written target item at a time and the lists were con-

structured so that each person would only see each item either with a left-dislocation or with a right-dislocation.³

6.

* La meua cosina s'ha canviat de pis i ahir hi va fer una festa.

Els ha pintat de color taronja, els radiadors.

1 2 3 4 5 6 7 8 9 10

Anterior Següent

De

SurveyMonkey

Vegeu fins a quin punt és fàcil [crear una enquesta.](#)

Figure 1. Example of target item

3.2 Predictions

Taking as a departing point the view represented Vallduví (1992), Villalba (2000: Chapter 3), and Bott (2007) that RD was better suited for closer anaphoric relationships, such as identity, we predicted that RD would be preferred in the most direct bridging types, namely those that required little or no deduction at all. Hence, EPITHET was clearly predicted to be better with RD than with LD, whereas at the other side of the scale, INDUCIBLE PARTS and OPTIONAL ROLE were predicted to be bad with RD, but fine with LD. As for less extreme cases, which involved a quite straightforward deduction step (HYPONYM, SET MEMBERSHIP, NECESSARY PARTS, and NECESSARY ROLE), we predicted LD to be preferred over RD by default.

4. Results

The 168 informants produced 7056 target answers. All analyses were performed with RStudio (RStudio 2014), running an R language (R Core Team 2014).

3. One anonymous reviewer pointed out to us that oral stimuli with controlled intonation contour should be preferred over potentially ambiguous written ones. We fully agree with his/her appreciation, and, indeed, we have included oral stimuli in a follow-up experiment in progress conducted with Lisa Brunetti.

4.1 Descriptive statistics

On average, informants rated RD higher (5.43) than LD (4.94), and they preferred RD over LD in all bridging types, with the exception of EPITHETS and NECESSARY ROLES, as can be appreciated in Figure 2.

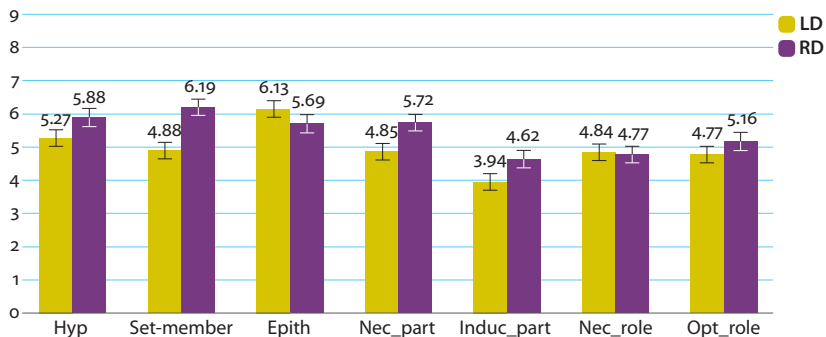


Figure 2. Means for LD and RD regarding bridging type

The normalized values (z-scores) are presented in Table 1, where the best ratings for RD are for HYPONYMS, SET MEMBERSHIP and NECESSARY PARTS, and LD only outscores RD clearly for EPITHETS and very slightly for NECESSARY ROLES.

Table 1. Normalized means by dislocation and bridging type

	Hyp	Set_member	Epith	Nec_part	Induc_part	Nec_Role	Opt_role
LD	0.025	-0.120	0.389	-0.146	-0.500	-0.136	-0.191
RD	0.295	0.425	0.170	0.208	-0.230	-0.177	-0.010

4.2 Significance tests

To test the significance of our results, we conducted t-tests for each bridging type regarding dislocation: at $p < 0.05$, the differences between LD and RD was significant for all bridging types, except NECESSARY ROLES; at $p < 0.001$, differences were significant for HYPONYM, SET MEMBERSHIP, NECESSARY PARTS and INDUCIBLE PARTS. Thus, RD is preferred for all bridging types, except EPITHETS, for which the preference is reversed, and NECESSARY ROLES, for which there is no preference.

When the related bridging types were compared, NECESSARY PARTS vs INDUCIBLE PARTS showed significant differences for both LD and RD, and NECESSARY ROLES vs OPTIONAL ROLES for RD only ($p < 0.001$). Overall, the types with closer bridging relationships (NECESSARY PARTS and NECESSARY ROLES) receive higher ratings than their counterparts with less tight relationships (INDUCIBLE PARTS and OPTIONAL ROLES).

To confirm the t-tests, we build a linear model with the z-score as answer variable and the type of dislocation and bridging as dependent variables. The differences were found to be highly significant ($p < 0.001$) for both LD and RD, and INDUCIBLE PARTS, NECESSARY PARTS, NECESSARY ROLES and OPTIONAL ROLES, and significant ($p < 0.01$) for HYPONYM and SET MEMBERSHIP. Finally, to control for individual variation, we build a linear mixed-effects model with the variable value as answer variable, the variable participant as a random factor and the bridging and dislocation types as independent variables. The model yielded highly significant effects ($p < 0.001$) for all bridging types, and a significant effect ($p < 0.05$) for dislocation type.

5. Discussion

The experiment shows that speakers consistently distinguish RD and LD uses, as the significance test show. Moreover, generally RD scores are significantly higher than LD scores overall, even in bridging cases typically suited for LD, namely those involving poset relations (HYPONYM and SET-MEMBERSHIP). Hence, Catalan seems to prefer RD as the unmarked dislocation mechanism, which confirms original claims by Laca (1986), Solà (1990, 1994) and Vallduví (1992) that RD was, among other things, a purely focalization mechanism which removes nonfocal material from the final sentence position where sentence stress and focus are assigned.

Regarding bridging types, according to the received view discussed in 2.2, we expected that the closer the bridging relation between the antecedent and the dislocate, the better RD would be. If we take EPITHET (*John-that idiot*) as the closest one, since it is a variant of identity relation, and INDUCIBLE PARTS (*kitchen-coffee maker*) and OPTIONAL ROLE (*death-murder*) as the more distant, we cannot confirm such an expectation, as we remarked in 4.1.

Just to reinforce the argument, consider the normalized means in Figure 3 (corresponding to the values in Table 1): LD only rates better than RD with EPITHET, and NECESSARY ROLE, which are under any point of view, closer bridging relations than INDUCIBLE PARTS or HYPONYM. Yet, these two later relations show a clear preference for RD.

Moreover, the role of poset relations with LD is not particularly clear, either. We have described in 2.2 that poset relations have been claimed to underlie the licensing of LD (Ward and Prince 1991; Vallduví 1992). As explained in Section 3, among the seven bridging types studied, three of them can be analyzed as poset relations: HYPONYM (*fruit-watermelon*), SET MEMBERSHIP (*Italian cities-Florence*), and EPITHET (*John-that idiot*). Yet, a contradictory result follows, for LD only rates best in the latter, epithet, which is a kind of identity relation, and hence better

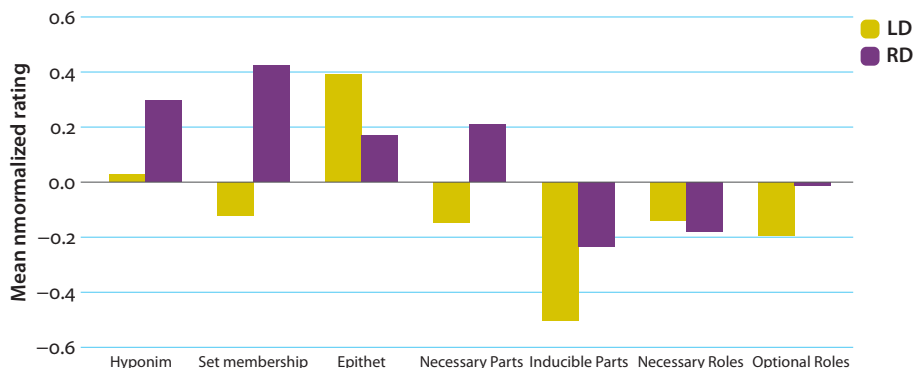


Figure 3. Normalized means for dislocation type across bridging types

suiting for RD. Moreover, the two other poset relations have a clear preference for RD. Overall, even though we didn't place poset relations as a condition in our experiment, the results cast doubts on the validity of the Discourse Condition on Preposing in Topicalization by Ward and Prince (1991: 173), which Vallduví (1992) endorses for Catalan LD.

We speculate that LD in Catalan may be subject to a requirement stricter than the Discourse Condition on Preposing in Topicalization, based on the poset relation. In other words, LD requires an overt contrast between the referent in the dislocated phrase and another referent. Thus, the prediction is that (10b), in which the watermelon is explicitly contrasted with the melon, would be higher rated than (10a), which was an item in our experiment and in which, although there is a poset relationship, there is no explicit contrast. We leave it for future work testing this contrast.

- (10) a. *Avui he menjat fruita de postres. La síndria, me l'he acabada tota.*
 'Today I've eaten fruit for dessert. The watermelon, I have eaten it all.'
- b. *Avui he menjat fruita de postres. La síndria, me l'he acabada tota; en canvi, el meló ni l'he tocat.*
 'Today I've eaten fruit for dessert. The watermelon, I have eaten it all; in contrast, the melon I have not even touched.'

6. Conclusions

In this article, we have shown that Catalan dislocation constructions fall into the realm of bridging phenomena, but in a very different way from what has been generally assumed in the literature. While it is clear that LD and RD are clearly perceived as different for all categories, the preferences are always on the RD side,

against common assumptions on the narrower distribution of RD: RD enters into all bridging relations more easily than LD, with the exception of EPITHET. Overall, the data show that Catalan speakers can use RD for a wider range of uses than assumed in previous works (Vallduví 1992, 1994, Vallduví and Engdahl 1996, Villalba 2000, 2009, Bott 2007). Henceforth, our study confirms from an experimental point of view the insights in Mayol (2007), Escandell-Vidal (2009), Villalba (2011), and Villalba and Mayol (2013), which call for a more flexible approach to the informative functions of RD.

Acknowledgements

We wish to thank the comments and suggestions by the audiences at UAB, and at LSRL (Stony Brook, 2016), and two anonymous reviewers, as well as funding provided by Ministerio de Ciencia e Innovación research projects FFI2011-23356, FFI2014-52015, FFI2015-66732-P, and FFI2015-67991-P.

References

- Asher, Nicholas, and Alex Lascarides. 1998. Bridging. *Journal of Semantics* 15 (1): 83–113. <https://doi.org/10.1093/jos/15.1.83>
- Birner, Betty J., and Gregory L. Ward. 1998. *Informational Status and Noncanonical Word Order*. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.40>
- Bott, Stefan. 2007. “Information Structure and Discourse Modelling.” Ph.D. dissertation, Universitat Pompeu Fabra.
- Brunetti, Lisa. 2009. “On Links and Tails in Italian.” *Lingua* 119 (5): 756–781. <https://doi.org/10.1016/j.lingua.2008.10.005>
- Clark, Herbert H. 1977. Bridging. In *Thinking: Readings in Cognitive Science*, ed. by Philip N. Johnson-Laird, and Peter C. Wason, 411–420. Cambridge: Cambridge University Press.
- Escandell-Vidal, Victoria. 2009. “Differential Object Marking and Topicality: The Case of Balearic Catalan.” *Studies in Language* 33 (4): 832–885. <https://doi.org/10.1075/sl.33.4.02esc>
- Grice, H. Paul. 1975. “Logic and Conversation.” In *Syntax and Semantics*, ed. by Peter Cole, and Jerry Morgan, 41–58. New York: Academic Press. <https://doi.org/10.1111/j.1365-2664.2006.01229.x>
- Grice, H. Paul. 1981. “Presupposition and Conversational Implicature.” In *Radical Pragmatics*, ed. by Peter Cole, 183–98. New York: Academic Press.
- Grice, H. Paul. 1989. *Studies in the Way of Words*. Cambridge, MA: Harvard University Press.
- Grosz, Barbara, and Yael Ziv. 1998. “Centering, Global Focus, and Right Dislocation.” In *Centering in Discourse*, ed. by Marilyn Walker, Aravind Joshi, and Ellen Prince, 39–51. Oxford: Oxford University Press.
- Hendriks, Herman. 1996. “Information Packaging: From Cards to Boxes.” In *Proceedings of Semantics and Linguistic Theory 6*, edited by Teresa Galloway, and Justin Spence, 6:75–92. Ithaca, NY: Cornell University.

- Hendriks, Herman, and Paul Dekker. 1995. "Links without Locations." In *Proceedings of the Tenth Amsterdam Colloquium*, ed. by Paul Dekker, and Martin Stokhof, 339–350. Amsterdam: ILLC.
- Irmer, Matthias. 2011. *Bridging Inferences. Constraining and Resolving Underspecification in Discourse Interpretation*. Berlin: De Gruyter. <https://doi.org/10.1515/9783110262018.223>
- Laca, Brenda. 1986. "Notes per a un estudi del pleonasme pronominal en català." In *Estudis de Llengua i Literatura Catalanes XIII. Miscel·lània Antoni Badia i Margarit 5*, ed. By Josep Massot i Muntaner, 65–88. Barcelona: Publicacions de l'Abadia de Montserrat.
- Lewis, David. 1979. "Scorekeeping in a Language Game." In *Semantics from Different Points of View*, ed. by Rainer Bäuerle, Urs Egli, and Arnim von Stechow, 339–359. Berlin: Springer-Verlag. https://doi.org/10.1007/978-3-642-67458-7_12
- López, Luis. 2009. *A Derivational Syntax for Informational Structure*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199557400.001.0001>
- Matsui, Tomoko. 1993. "Bridging reference and the notions of *topic* and *focus*." *Lingua* 90 (1–2): 49–68. [https://doi.org/10.1016/0024-3841\(93\)90060-A](https://doi.org/10.1016/0024-3841(93)90060-A)
- Mayol, Laia. 2007. "Right-dislocation in Catalan Its discourse function and counterparts in English." *Languages in Contrast* 2 (1): 203–219. <https://doi.org/10.1075/lic.7.2.07may>
- R Core Team. 2014. *R: A language and environment for statistical computing* (Version 3.1.2, 64 bits) [Computer software]. Vienna: R Foundation for Statistical Computing. <http://www.R-project.org/>.
- RStudio. 2014. *RStudio: Integrated development environment for R* (Version 0.98.1091) [Computer software]. Boston: RStudio. <http://www.rstudio.org/>
- Simons, Mandy. 2003. "Presupposition and Accommodation: Understanding the Stalnakerian Picture." *Philosophical Studies* 112 (3): 251–78. <https://doi.org/10.1023/A:1023004203043>
- Solà, Joan. 1990. *Lingüística i normativa*. Barcelona: Empúries.
- Solà, Joan. 1994. "Ordre d'elements I duplicats pronominals." In *Sintaxi normativa. Estat de la qüestió*, 235–63. Barcelona: Empúries.
- Stalnaker, Robert C. 2002. "Common Ground." *Linguistics and Philosophy* 25: 701–21. <https://doi.org/10.1023/A:1020867916902>
- Stalnaker, Robert C. 2014. *Context*. Oxford: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199645169.001.0001>
- Vallduví, Enric. 1992. *The Informational Component*. New York: Garland Publishing.
- Vallduví, Enric. 1994. "Detachment in Catalan and Information Packaging." *Journal of Pragmatics* 22 (6): 573–601. [https://doi.org/10.1016/0378-2166\(94\)90031-0](https://doi.org/10.1016/0378-2166(94)90031-0)
- Vallduví, Enric, and Elisabet Engdahl. 1996. "The Linguistic Realisation of Information Packaging." *Linguistics* 34 (3): 459–519. <https://doi.org/10.1515/ling-2013-0041>
- Villalba, Xavier. 2000. "The Syntax of Sentence Periphery." Ph.D. dissertation, Universitat Autònoma de Barcelona.
- Villalba, Xavier. 2009. *The Syntax and Semantics of Dislocations in Catalan. A Study on Asymmetric Syntax at the Peripheries of Sentence*. Köln: Lambert.
- Villalba, Xavier. 2011. "A Quantitative Comparative Study of Right-Dislocation in Catalan and Spanish." *Journal of Pragmatics* 43 (7): 1946–1961. <https://doi.org/10.1016/j.pragma.2011.01.002>
- Villalba, Xavier, and Laia Mayol. 2013. "Right-Dislocation in Catalan: Tails, Polarity and Activation." *International Review of Pragmatics* 5 (1): 87–117. <https://doi.org/10.1163/18773109-13050104>

- Wall, Edward. 1972. *Introduction to Mathematical Linguistics*. Englewood Cliffs, NJ: Prentice-Hall.
- Ward, Gregory L. 1988. “The Semantics and Pragmatics of Preposing.” Ph.D. Dissertation, University of Pennsylvania.
- Ward, Gregory L., and Ellen F. Prince. 1991. “On the Topicalization of Indefinite NPs.” *Journal of Pragmatics* 15 (2): 167–178. [https://doi.org/10.1016/0378-2166\(91\)90079-D](https://doi.org/10.1016/0378-2166(91)90079-D)
- Ziv, Yael. 1994. “Left and Right Dislocations: Discourse Functions and Anaphora.” *Journal of Pragmatics* 22 (6): 629–645. [https://doi.org/10.1016/0378-2166\(94\)90033-7](https://doi.org/10.1016/0378-2166(94)90033-7)

Dependent numerals and dependent existentials in Romanian

Mara Panaitescu

University of Bucharest

The paper discusses Romanian dependent indefinites (DepIndef's) introduced by the distributive marker *câte* (/ˈki.te/). They come in two guises, referred to as dependent numerals (DepNum's) in the configuration [CÂTE NUM NP], and dependent existentials (DepExist's), in the configuration [CÂTE SG-INDEF NP]. The two marked indefinites are identical to their unmarked counterparts at the level of assertion. The narrower distribution of DepNum's is an effect of the interaction between the numeral and certain Partition operators which license *câte*. The licensing condition of *câte* DPs in both guises states that these DPs must be under the scope of a D/Part operator in the terminology of Champollion (2016a,b). DepNum's are associated with an event-dependent measure of cardinality.

Keywords: dependent indefinites, distributive operator, partition

1. Introduction

Teasing apart DepExist's and DepNum's is complicated because the Romanian singular indefinite, *un_{masc} / o_{fem}*, is also the singular cardinal. Nevertheless, the differences in distribution between singular and plural forms indicate that: (i) when the singular form is perfectly acceptable, but a DepNum with a cardinal higher than *one* is not, I assume (with Farkas 2015) that the singular form is the singular existential indefinite rather than the singular cardinal; (ii) it follows that DepExist's have a wider distribution than DepNum's. To take an example from Brasoveanu and Farkas (2011), in (1), where only the singular is allowed, *câte o gară* is considered an instance of the DepExist ('CÂTE a station') rather than the DepNum ('CÂTE one station'):

- (1) *Din când în când, trenul se oprea în câte o gară/ *în câte două gări.*
 from when in when train.DEF REFL stopped.IMPV in CÂTE a station in CÂTE
două gări.
 two stations
 ‘Occasionally, the train stopped in a station/ *in two stations.’

Moreover, the paper is dedicated to adnominal uses of *câte* – where by adnominal I mean “syntactically part of a DP” – of the form [CÂTE NUM NP] or [CÂTE SG-INDEF NP]. At a certain point in the argumentation, a different kind of *câte*-phrase will be illustrated, which will be referred to as adverbial *câte*, in the form [NUM CÂTE NUM]. The distribution of this construction indicates that adverbial *câte*, but not adnominal *câte*, has a semantics close to that of English “Num by Num” constructions (*one by one, two by two etc.*).

The following section addresses two important questions in the literature on dependent indefinites. Firstly: whether DepIndef’s come with their own distributivity operator. Section 2.1 provides evidence to the contrary, in the case of DepNum’s and DepExist’s respectively. Secondly: what kind of plurality may serve as sortal key for DepIndef’s? It will be shown that, for DepNum’s and DepExist’s alike, the key may be a participant or a spatio-temporal dimension.

2. The distribution of DepIndef’s in Romanian

2.1 Does “*câte*” supply distributivity?

Regarding the question whether distributive force is supplied by *câte* itself, the data suggest that the answer is no. The evidence comes from sentences *fiecare* (‘each’), which already provides a distributive operator. Unlike English sentence %Each student read a book each, which is only acceptable for certain speakers,¹ (2) below is perfectly grammatical.

- (2) *Fiecare copil a cumpărat câte două baloane.*
 each child has bought CÂTE two balloons
 ‘Each child bought two balloons (per child).’

Above, the sortal key is the subject DP and the distributivity provided by the quantifier *fiecare* is sufficient to license the DepNum with no redundancy.²

1. Szabolcsi (2010, 2011).

2. Actually, (c) is preferable to a sentence containing only adnominal *fiecare* and no *câte*. See also Dotlacil (2015) for experimental data on two Czech distributive markers, corresponding to

Further evidence comes from sentences with two DepNum's but only one Dist (interpretation (a) below):

- (3) *Băieții au cumpărat câte două cadouri la câte trei fete.*
 boys_{DEF} have bought CÂTE two gifts to CÂTE three girls
 'The boys bought two gifts to three girls.'

Assuming there are four boys: these are the possible interpretations:

- (3') a. 8 gifts, 12 girls – gifts and girls both vary with boys
 b. 8 gifts, 24 girls – gifts vary with boys, but girls vary with gifts
 c. 24 gifts, 12 girls – gifts vary with girls and girls vary with boys

(3'a) indicates that a single distributive key (*the boys*) can validate more than one DepNum. If DepNum's contribute distributivity, the distributive force cannot have its source in both DepNum's (in that case, (3'b,c) would be the only attested readings).

Examples of inverse scope *câte* seem to go against this conclusion. As observed in Zimmermann (2002) and Champollion (2016b), German distributive marker *jeweils* may occur in subject position, with the direct object apparently playing the role of sortal key.

- (4) a. *Jeweils ein Offizier begleitete die Ballerinen nach Hause.*
 DIST one officer accompanied the ballerinas to home
 'The ballerinas were accompanied home by one officer each.'
 b. *Câte un ofițer a acompaniat balerinele acasă.*

The Romanian translation of (4), given in (b), is also acceptable (as well as versions with higher cardinals). Yet, such inverse scope sentences are in the general case difficult to interpret in Romanian. (4) is only interpreted in a setting involving "organized events". Compare (4) with (5), which is degraded because it is difficult to understand the seeing and killing events as part of a larger organized event:

- (5) ^{??}*Câte doi ofițeri au văzut/ omorât balerinele.*
 CÂTE two officers have seen/ killed ballerinas.def
 Intended: 'Officers in pairs have seen/ killed the ballerinas.'

Moreover, if the right contextual setting of "organized events" is present, a plural direct object is not required for distributivity. Assuming the soldiers are supposed to guard the treasure in shifts:

câte (po) and *fiecare (každy)*, which display the same effect: reading times decrease if both are present, as opposed to only adnominal *každy*.

- (6) *Câte doi soldați vor veghea comoara până mâine dimineață.*
 CÂTE two soldiers will guard treasure.def until tomorrow morning
 ‘Two soldiers will guard the treasure until tomorrow morning (in shifts).’

We conclude that DepNum’s in inverse scope do not introduce themselves, but require the presence of a contextually provided partition of events, where the sortal key is a contextually provided runtime (e.g. ‘shifts’ in the example above). If no partition is retrievable, the sentence is uninterpretable.

2.2 What type of variable may serve as key?

2.2.1 *DepNums*

Farkas (2015), discussing Hungarian reduplicated numerals, proposes that DepNum’s, unlike DepExist’s may only be licensed by individual variables (not by event variables). Evidence provided includes iterative adverbs and pluractionals, which don’t license DepNum’s in Hungarian. These facts also apply to Romanian.

Henderson (2016), discussing the same data, proposes that all DepNums are licensed due to the existence of a plural event variable; DepNum’s in general are not selective with respect to the type of variable that licenses them. DepNum’s in Hungarian differ from DepExist’s in Hungarian and DepNum’s in Kaqchikel, in that they also require to be under the scope of a distributive operator. Thus, iterative adverbs and pluractional verbal markers, which meet the event plurality condition, but do not supply a distributive operator, can license only the latter.

In the case of Romanian DepNum’s, the evidence suggests that *câte* is not only licensed by individual variables – (6) above and (7) below, a modified variant from Farkas (2015). Suppose someone is reporting what happened in a children’s ward in the course of one night:

- (7) a. *??S-au mai trezit (câte) doi copii, dar in rest nimic.*
 REFL-have MAI woken CÂTE two children but in rest nothing
 ‘Two children woke up (from time to time), but nothing besides that.’
- b. *S-au mai trezit cel mult (câte trei copii/ (câte) doi sau trei*
 REFL-have MAI woken the most CÂTE three children CÂTE two or three
copii/ (câte) doi copii deodată.
 children/ CÂTE two children simultaneously
 ‘At most three children/ two or three children/ two children
 simultaneously woke up (from time to time).’

The Romanian adverb *mai* contributes an interpretation of the sentence which can be informally stated as ‘there were some/ a few random occasions such that...’. It was argued in Section 2.1 that *câte* cannot itself introduce distributivity, yielding ‘there were some/ a few occasions such that on each of these occasions two

children woke up”. Thus, whatever property of these iterated adverbs is responsible for the unacceptability of non-specific readings of plain cardinals in (7) (with pairs of children varying with the relevant occasions introduced by *mai*) is also involved in the unacceptability of DepNum’s. More precisely, *câte* in (7) does verify its licensing condition, but the sentence is unacceptable due to the presence of the unmodified cardinal. Whenever the plain cardinal on its non-specific reading becomes acceptable, as in (7), the DepNum is also acceptable. A DepExist does not run into the same problem, hence the wider distribution. Regarding Henderson’s distinction between DepNum’s and DepExist’s, it follows that no extra constraint on DepNum’s is required. If the presence/ absence of a distributive operator were the relevant factor, the acceptability of modified cardinals and modified DepNum’s in (7) would remain unexplained. Both DepNum’s and DepExist’s will be analyzed as being subject to the same licensing condition: they must be in the scope of a D/ Part operator (see Section 3).

2.2.2 *DepExist’s*

The contrast between plain existential indefinites and DepExist’s seems to be weaker than assumed in Farkas (2015), at least inasmuch as the Romanian data suggest. The examples in Farkas placed the DepExist’s in topic position. In Hungarian, this position allows for the co-variation of individual children with situations only when the singular indefinite is reduplicated:

- (8) a. *Egy-egy gyerek fel-felébredt de más baj nem volt.*
 a-a child up-up woke, but other trouble was not.
 ‘Repeatedly, some child or other woke up.’
 b. *Egy gyerek fel-felébredt de más baj nem volt.*
 a child up-up woke, but other trouble was not.
 ‘Repeatedly, a child (the same one) woke up.’

In Romanian, the plain indefinite is not so well-behaved. For instance it does seem to get co-varying readings with *mai*, especially in inner VP position (9a), but not with iterative adverb *tot*, which indicates an iteration with a high frequency (translated with the English *keep V-ing* construction in (9b)).

- (9) a. *S-a mai trezit (câte) un copil.*
 REFL-has MAI woken CÂTE a child
 ‘From time to time, a child woke up.’
 a.i. without CÂTE: the same/ different children
 a.ii. with CÂTE: different children
 b. *S-a tot trezit (câte) un copil.*
 REFL-has TOT woken CÂTE a child
 ‘A child kept waking up.’

- b.i. without CÂTE: the same child
- b.ii. with CÂTE: different children

The contrast between narrow-scoping indefinites and *câte*-marked indefinites is therefore taken to be due to three cumulated effects on plain cardinals/ existentials: the topic position, the availability of a competing specific reading and the licenser itself. DepNum's and DepExist's with event distributive readings need a (contextually provided) partition, plain cardinals and existentials do not. I leave the explanation for the differences between licensors to future inquiry (not only *tot* and *mai*, but also frequency adverbs behave differently among themselves with respect to licensing potential). The conclusion is that neither DepNum's, nor DepExist's are restricted to participant readings i.e. the distributive key need not be an individual variable. Furthermore, the comparison between Romanian plain cardinals and DepNum's suggested that the two contribute the same semantic content at the level of assertion.

3. The licensing conditions of *câte* dependent indefinites

3.1 Champollion's D and Part operators

Champollion (2016a,b) provides a unified account of phrasal distributivity. All instances of phrasal distributivity have in common the presence of a distributivity operator. The different kinds of distributivity (participant-based vs. event-based) and the manner of dividing the sortal key (into atoms or groups) are decided by two properties of the relevant operator, which comes in two guises (D and Part).³ The two (related) properties are granularity and dimension. Considering granularity, D is set to atoms only, while Part is a cover-based operator which can also distribute over non-atomic pluralities under contextual licensing. Considering dimension, an operator may establish a dependency with respect to a thematic role or a temporal/ spatial dimension of events (that is, the key may be a participant or a runtime/ spatial coordinate). The two properties (granularity and dimension) are related: setting the granularity parameter to non-atomic in the theta-role dimension is difficult to obtain and requires contextual support; setting the granularity parameter to atomic in the temporal/ spatial dimension is impossible. Thus,

3. More precisely, D is a subcase of Part. An example where this fact becomes evident is Champollion's entry for the German distributive marker *jeweils*, which introduces Part, and whose distribution includes the theta-role dimension with granularity atom, in which cases it becomes synonymous to English binominal *each*.

we expect to find D naturally and Part with certain difficulty for participant-based distributivity; and we expect only Part to apply to the spatial/temporal dimension.

Firstly, D and Part are assumed to have both covert and overt instantiations. Secondly, they are taken to be indexed with the theta-role/ runtime function of the sortal key. Thirdly, the two operators are formulated in Neo-Davidsonian event semantics, such that they apply to predicates of events. Concerning this, irrespective of whether they induce a participant-based or an event-based kind of distributivity, the D/Part operators always target events via some property of events, theta role or runtime.

- (10) *Definition: Event-based D operator*
 $[[D_\theta]] = \lambda V \lambda e [e \in * \lambda e' (V(e') \ \& \ \text{Atom}(\theta(e')))]$

D is indexed with a theta variable (the key) and requires the relevant subevents e' to have atoms as key participants; the definition above further requires the larger event e to be an element of the set containing subevents e' (and sums thereof) whose participant corresponding to the indexed theta role is atomic. See below for an example:

- (11) a. *The boys* [D_{agent} [*saw a monkey*]].
 b. $\exists e [* \text{agent}(e) = \oplus \text{boy}$
 $\ \& \ e \in * \lambda e' (* \text{see}(e') \ \& \ \text{monkey}(\text{theme}(e')) \ \& \ \text{Atom}(\text{agent}(e')))]$

There is an event e whose Agent is the boys and this event is an element of a set containing events e' of seeing in which a monkey is the theme and a boy is the agent, as well as the sums of such events, as indicated by the 'star' operator. Of course, e can only be a sum of such events, namely the sum whose agent is the boys. As illustrated below, D may also target some other theta role:

- (12) a. *The first-year students* [D [*took an exam*]]. Target: agent
 b. *John* [D [*gave a pumpkin pie*]] *to two girls*. Target: goal

Although Champollion does not mention this, double distributive readings are expected to be derived by the introduction of two D's, each co-indexed with a different theta role, as in (13). Thus, the application of D is expected to be recursive. We assume the representation (13).⁴

- (13) a. *The first-year students* [D_{agent} [D_{goal} [*gave a pumpkin pie*]] *to two girls*]].
 b. $\exists e [* \text{agent}(e) = \oplus \text{f-y-students} \ \&$
 $\ e \in * \lambda e' (* \text{give}(e') \ \& \ * \text{girl}(\text{goal}(e')) \ \& \ |\text{goal}(e')| = 2 \ \& \ \text{Atom}(\text{agent}(e')) \ \&$
 $\ e' \in * \lambda e'' (\text{give}(e'') \ \& \ \text{pumpkin-pie}(\text{theme}(e'')) \ \& \ \text{Atom}(\text{goal}(e'')))]$

4. The same mechanism applies in reading (3).

(The students are the agents of giving events e such that in all subevents e' of e , each of the students gave something to two girls and in each subevent of e' , each girl received a pumpkin pie)

Two properties of events and theta-roles (namely cumulativity and sum homomorphism, see Champollion's 2016a definitions (34) and (35)) ensure that the subevents 'keep track' of their theta role. That is, looking at (13), since the agent of e is *the students*, e can only be a sum of events e' with atomic agents which are necessarily also students. Conversely, if e' is a sum of events made up of events e'' that have a pumpkin pie as theme, e' can only have (a sum of) pumpkin pies as theme and nothing else.

The distributive operator Part differs from D in that its granularity is not set to atoms.

- (14) *Definition: Event-based Part operator*
 $[[\text{Part}_{\theta,C}]] = \lambda V \lambda e [e \in * \lambda e' (V(e') \ \& \ C(\theta(e')))]$

C is a set and θ is either a theta role or a temporal/ spatial property of an event. $C(\theta(e'))$ means that $(\theta(e'))$ is an element of C . An example where C ranges over theta roles:

- (15) a. *The shoes* $[\text{Part}_{\text{theme,pair}} [\text{cost fifty dollars}]]$.
 b. $\exists e [* \text{theme}(e) = \oplus \text{shoes} \ \& \ \text{Part}_{\text{theme,pair}} ([[\text{cost fifty dollars}]](e))]$
 $\Leftrightarrow \exists e [* \text{theme}(e) = \oplus \text{shoes} \ \& \ e \in * \lambda e' (e' \in [[\text{cost fifty dollars}]] \ \& \ \text{pair}(\text{theme}(e')))]$

C is contextually defined as a set of pairs and co-indexed with the theme theta role. (15) says that there is a plural event whose themes sum up to the shoes and which consists of costing-fifty-dollars events with pairs as themes. Part with non-atomic granularity in the theta role dimension is limited.

An example where multiple runtimes are involved but distributivity can only occur if a contextual cover is provided are sentences with *for*-adverbials (Champollion 2016b, Example (73), represented as (16) below). Although *for*-adverbials do not generally distribute over indefinite direct objects, (16) below, uttered in the specified context, means that the patient took two different pills with a certain regularity within the interval of a month.

- (16) Context: discussing daily pill intake
The patient took two pills for a month (and then went back to one pill).

Assuming the contextual partition has cells of the size of one day (the patient took two pills every day for a month), the partition operator will be $\text{Part}_{\tau, \lambda t. \text{days}(t) \leq 1}$, i.e. a partition which is set for the dimension 'runtime' and whose contextual cover

is set for intervals shorter than or equal to one day. For-adverbials are assumed to come with a test verifying whether the verb is regular and atelic; if it meets this condition, the *for*-adverbial passes the denotation of the verb phrase up unchanged. The result of applying Part and the *for*-adverbial is:

- (17) $\exists e : \text{regular}(\tau(e)). *agent(e) = \text{the.patient} \ \& \ \text{months}(\tau(e)) = 1 \ \& \ e \in * \lambda e'$
 $(* \text{take}(e') \ \& \ * \text{pill}(* \text{theme}(e')) \ \& \ | * \text{theme}(e') = 2| \ \& \ \text{days}(\tau(e')) \leq 1]]$
(There is a regular plural event that consists of one or more events of taking two pills which each take place within a day. Its agent is the patient, and its runtime measures a month.)

Having presented the covert instances of D and Part, we now move on to overt instantiations. Champollion (2016b) assumes that distributive markers contain a D or Part operator in their denotation. English *each*, which may have thematic roles but not runtimes as distributive key and whose granularity is restricted to Atom, is assumed to be an instantiation of D. German *jeweils*,⁵ which applies to both theta roles and runtimes and whose granularity is not restricted to Atom, is assumed to be an instantiation of Part.

- (18) *Adnominal each*
 $[[\text{each}_{\theta}]]_{\text{adnominal}} = \lambda P \lambda \Theta \lambda e. e \in [[D_{\theta}]](\lambda e'. P(\Theta(e')))$
 where D_{θ} = definition (10)

In the entry above, θ stands for the thematic role of the key, while Θ is the thematic role of the share. For example, *The boys $_{\theta}$ saw two monkeys $_{\Theta}$ each* is true of an event e whose Agent (θ) is the boys and which is made up of events e' with atomic agents and two monkeys as theme (Θ).

- (19) *Adnominal jeweils*
 $[[\text{jeweils}_{\theta,C}]]_{\text{adnominal}} = \lambda P \lambda \Theta \lambda e. [[[\text{Part}_{\theta,C}]](\lambda e'. [P(\Theta(e'))]) (e)$
 $\ \& \ (C \neq \text{atom} \rightarrow \oplus C = \theta(e))]]$
 $= \lambda P \lambda \Theta \lambda e. e \in * \lambda e' [P(\Theta(e')) \ \& \ C(\theta(e'))]$
 $\ \& \ (C \neq \text{atom} \rightarrow \oplus C = \theta(e))]]$

In the sentence below, θ is set to Agent and C is set to atoms:

- (20) *Die Jungen haben jeweils $_{\theta=\text{agent},C=\text{atom}}$ zwei Affen gesehen.*
 the boys have DIST two monkeys seen
 'The boys have seen two monkeys each.'

This setting yields the same interpretation as that of adnominal *each* in *The boys saw two monkeys each*. Unlike *each*, however, *jeweils* also has event-based

5. See Zimmermann (2002) for the distribution of adnominal *jeweils*.

distributive readings (more precisely runtime-based), if the context makes this reading salient. For example, if the children went on zoo visits multiple times, (21) below can be uttered to express that, on each zoo visit, they saw two monkeys.

(21) *Die Jungen haben jeweils* _{$\theta=\tau, C=zoo\text{-visit}$} *zwei Affen* _{$\Theta$} *gesehen*
 the boys have DIST two monkeys seen
 ‘The boys have seen two monkeys each time.’

(22) $\exists e [*agent(e) = \oplus boy \ \& \ *see(e) \ \& \ e \in * \lambda e' [|(e') = 2| \ \& \ *monkey(theme(e')) \ \& \ zoo\text{-visit}(\tau(e'))] \ \& \ \oplus zoo\text{-visit} = \tau(e)]$

3.2 Romanian DepNum and DepExist

The discussion in Section 2 has yielded the following results: dependent indefinites (both cardinal and existential) do not supply a distributive/ partition operator themselves. These indefinites must be in the scope of some such operator. The contrast between adnominal and adverbial uses of *câte* suggests that only the adverbial version comes with its own partition, hence the incompatibility with a distributive key in (24), which is marked # on the interpretation in which the distribution of cakes per boy is two – an intended counterpart of (24).

(23) *Definite plural key*

a. *adverbial câte*

Copiii au luat prăjituri, două câte. două de pe platou
 children.DEF have taken cakes two CÂTE two from on platter.DEF
 ‘The children took cakes in two by two from the platter.’

b. *adnominal câte*

Copiii au luat câte două prăjituri de pe platou.
 children.DEF have taken CÂTE two cakes from on platter.def
 ‘The children took two cakes per child from the platter.’

(24) *Universal distributive Q as key*

a. *adverbial câte*

(#)*Fiecare copil a luat prăjituri două câte două de pe platou.*
 each child has taken cakes two CÂTE two from on platter.def
 ‘Each child took cakes two by two from the platter.’

b. *adnominal câte*

Fiecare copil a luat câte două prăjituri de pe platou.
 each child has taken CÂTE two cakes from on platter.def
 ‘Each child took two cakes per child from the platter.’

This suggests adverbial *câte* comes with its own partition and could be analyzed in terms of encapsulation in the terminology of Brasoveanu and Henderson (2009),

on a par with English *one by one*. On the other hand, DepIndef's introduced by *câte* do not themselves supply D/Part, but must be in the scope of D/Part. The fact that both D and Part comply with the licensing conditions of DepIndef's is supported by the existence of contexts in which the granularity of the sortal key is not set to atoms (where the sortal key can be either a thematic role (a) or a runtime (b,c)):

- (25) a. *Pantofii costă câte 100 de lei.*
 shoes.DEF COST CÂTE 100 of lei
 'The shoes cost 100 lei per pair.'
- b. *Citesc câte două articole săptămânal.*
 read.I.SG CÂTE two articles weekly
 'I read two articles weekly.'
- c. Context: discussing daily pill intake
Pacientul a luat câte două pastile timp de o lună
 patient.DEF has taken CÂTE two pills time of a month
 'The patient took two pills for a month.'

3.2.1 The semantic contribution of "câte" in DepIndef's

The following licensing condition is formulated for DepIndef's introduced by *câte* (applying to both DepNum and DepExist). In the case of DepNum's, the measure of cardinality applies not to the individual variable introduced by the *câte*-DP, but to event-individual pairs in the sense of Doetjes and Honcoop (1997). The outcome is that DepNum's are interpreted as necessarily distributive and event-related DPs, that is, as the distributive counterpart of cumulative event-related numerals analyzed in Krifka (1990) and Doetjes and Honcoop (1997). See Panaitescu (to appear) for a detailed argumentation in favor of this treatment.⁶

(26) Evaluation condition:

For any variable x introduced by *câte* s.t. $P(x) \ \& \ \Theta(e,x)$, there is a D/Part such that:

a. DepNum

$D_{\theta} / \text{Part}_{\theta,C} (\dots [[\text{CÂTEN } P]]_{\Theta} \dots)$, where $\theta \neq \Theta$;

- (i) $D_{\theta} ([[\text{CÂTEN } P]])$: $\lambda V \lambda e. e \in * \lambda e' (V(e') \ \& \ \text{Atom}(\theta(e')) \ \& \ *P(\Theta(e')) \ \& \ |e', \Theta(e')| = n \dots)$
- (ii) $\text{Part}_{\theta,C} ([[\text{CÂTEN } P]])$: $\lambda V \lambda e. e \in * \lambda e' (V(e') \ \& \ C(\theta(e')) \ \& \ *P(\Theta(e')) \ \& \ |e', \Theta(e')| = n \dots)$

6. The only mentionable distinction between the account developed here and that in Panaitescu (to appear) is that the distributivity constraint is understood as a kind of polarity dependency here and as anaphoric dependency in the cited paper. I leave the advantages of a choice of one strategy over the other open for now.

b. **DepExist**

$$D_{\theta} / \text{Part}_{\theta,C} (\dots [[\text{C}\hat{\text{A}}\text{T}\text{E } a P]]_{\Theta} \dots), \text{ where } \theta \neq \Theta$$

- (i) $D_{\theta} ([[\text{C}\hat{\text{A}}\text{T}\text{E } a P]]): \lambda V \lambda e. e \in * \lambda e' (V(e') \ \& \ \text{Atom}(\theta(e')) \ \& \ P(\Theta(e'))$
 (ii) $\text{Part}_{\theta,C} ([[\text{C}\hat{\text{A}}\text{T}\text{E } a P]]): \lambda V \lambda e. e \in * \lambda e' (V(e') \ \& \ C(\theta(e')) \ \& \ P(\Theta(e')) \dots)$

In words: the thematic role Θ of the **DepNum** ($[[\text{c}\hat{\text{a}}\text{t}\text{e } n P]]$) and the **DepExist** ($[[\text{c}\hat{\text{a}}\text{t}\text{e } a P]]$) must occur in the scope of either a D operator (subconditions (a(i)) and (b(i))) or Part (subconditions (a(ii)) and (b(ii))), which is independently provided by an argument with a different thematic role or by a runtime. The formulation follows Champollion's (2016b) account of binominal *each* and *jewels*, but the difference resides in the fact that the distributive operator is not part of the denotation of the **DepIndef**. In this respect, this condition follows the same lines as and Oh (2001, 2006), who analyzes **DepIndef**'s as distributive-polarity items. The choice of talking in terms of partitions instead of dependent variables as in Brasoveanu and Farkas (2011) is motivated by the dependency of **DepIndef**'s on runtimes, which are functions over events. Runtimes are intervals which are included in a temporal variable t provided by Tense, but are not themselves introduced as variables in the derivation. It can be easily seen that the dependency induced by both D and Part is strictly speaking event dependency (meaning that participant distributivity is just a kind of event distributivity in which the subevents are individuated via some other participant – the key).

The measure of cardinality for **DNum**'s applies not to objects directly, but to event-object pairs, which corresponds to the intuition that the individuals in the share are only “possibly different”. More importantly, this explains the requirement that *c\hat{a}t\text{e}* should only apply to monotonic measures first noted in Zhang (2013) for binominal *each*. Cardinality is monotonic on nouns by default, so is weight when ranging over nouns denoting individuals that have mass, but temperature isn't. Since the measures of event-object pairs disregard the identity of the objects as long as the events are distinct, they obey additivity. For instance, the sum of three event-object pairs with events e_1 , e_2 and e_3 and participants of cardinality 2, will have cardinality 6 irrespective of whether the same individual shows up twice as participant in one of the three events, as long as the events are distinct. Since applying event-object pairs to non-monotonic dimensions leads to contradiction, *c\hat{a}t\text{e}* applied on measures of temperature in (27) is deviant:

- (27) *Cele dou\text{a} obiecte au c\hat{a}t\text{e} 10 kg/ \#20 de grade Celsius.*
 those two objects have $\text{C}\hat{\text{A}}\text{T}\text{E}$ 10 kg 20 of degrees Celsius
 ‘The two objects weigh ten kilos each/ measure 20 degrees Celsius each.’

Assume there are two liquids, a and b, in separate containers and they separately reached 20 degrees Celsius. Under the present account, a has 20 degrees and b has

20 degrees, a and b together have 20 degrees, and not 40 degrees, as the cardinality function in (26) would predict.

3.2.2 *Universally quantified sentences*

Champollion (2016b) analyzes the distributive quantifier *each* as introducing a distributive operator D, like its adverbial and adnominal counterparts:

$$(28) \quad \begin{aligned} [[\text{each}]_{\text{determiner}}] &= \lambda P \lambda \theta \lambda V \lambda e [\theta(e) = \oplus P \ \& \ [[D_{\theta}]](V)(e)] \\ &= \lambda P \lambda \theta \lambda V \lambda e [\theta(e) = \oplus P \ \& \ e \in * \lambda e' [V(e') \\ &\quad \& \ \text{Atom}(\text{agent}(e'))]] \end{aligned}$$

I assume the same entry for Romanian *fiecare*. In a sentence containing a DepInDef, the distributive operator D introduced by *each* takes scope over the indefinite, thus satisfying the evaluation condition. *Câte* further verifies the cardinality condition – the total lifted suitcases can be calculated as the number of boys times two:

- (29) a. *Fiecare băiat a ridicat câte două valize.*
 each boy has lifted CÂTE two suitcases
 b. $\exists e [\text{agent}(e) = \oplus \text{boy} \ \& \ e \in * \lambda e' [* \text{carry}(e') \ \& \ \text{Atom}(\theta(e')) \ \& \ |e', * \text{theme}(e')| = 3 \ \& \ * \text{suitcase}(* \text{theme}(e'))]]$

3.2.3 *Temporal keys*

Turning now to the cases where the key is a temporal partition (in the present case into shifts), the subject participant is within the scope of this partition and is interpreted as the share, thus it satisfies the conditions imposed by *câte*.

- (30) a. *Câte doi soldați vor veghea comoara până mâine-dimineată.*
 CÂTE two soldiers will guard treasure.DEF until tomorrow-morning
 ‘Two soldiers will guard the treasure until tomorrow morning (in shifts)’
 b. $\exists e: \text{regular}(e). \tau(e) < \text{tomorrow-morning} \ \& \ e \in * \lambda e' (* \text{guard}(e') \ \& \ * \text{soldier}(* \text{agent}(e')) \ \& \ |e', * \text{agent}(e')| = 2 \ \& \ * \text{theme}(e') = \text{the.treasure} \ \& \ \text{shift}(\tau(e'))$

In (b) above, it is implicitly assumed that the contextual Part operator imposes the condition *regular* (introduced in Vlach 1993, see Champollion 2016a for discussion in the context of *for*-adverbials) which, depending on the contexts, regulates the amount of discontinuity allowed. In the case at hand, no discontinuity is allowed and the shifts must be in strict succession (compare: *The patient will take two pills till next Tuesday*).

4. Conclusion

The previous sections provide new data from Romanian, which supports the view that DepIndef's introduced by *câte* do not supply a distributive operator themselves, unlike their adverbial counterpart (see (24)). Also, cardinality condition in (26) groups together DepNum's and *câte* in amount DPs. The data presented here invites further investigations regarding the interpretation of adverbial *câte* and amount *câte*, as well as a more detailed account of the interaction between cardinals and *for*-adverbials, pluractional markers and frequency adverbs. One general outcome of the discussion is that *câte* is always event dependent: even when D is involved, it's a plurality of events that licenses this distributive marker. That plurality in turn is individuated via atomic key participants.

Acknowledgements

This work is part of the program "Investissements d'Avenir, overseen by the French national Research Agency, ANR-10-LABX-0083. It is based on the presentation delivered together with Carmen Dobrovie Sorin at LSRL 46. Many thanks to Carmen Dobrovie Sorin for the advice, the insights and the critical comments, without which this paper could never have been written and to Lucas Champollion for helpful comments and clarifications.

References

- Brasoveanu, Adrian, and Donka F. Farkas. 2011. "How Indefinites Choose Their Scope." *Linguistics and Philosophy* 34 (1): 1–55. <https://doi.org/10.1007/s10988-011-9092-7>
- Brasoveanu, Adrian, and Robert Henderson. 2009. "Varieties of distributivity: One by one vs. each." *Proceedings of SALT 19*, ed. by Ed Cormany, Satoshi Ito, and David Lutz, 55–72. Ithaca, NY: Cornell University.
- Champollion, Lucas. 2016a. "Overt Distributivity in Algebraic Event Semantics." *Semantics and Pragmatics* 9 (16). 1–65. <https://doi.org/10.3765/sp.9.16>
- Champollion, Lucas. 2016b. "Covert Distributivity in Algebraic Event Semantics." *Semantics and Pragmatics* 9 (15). 1–66. <https://doi.org/10.3765/sp.9.15>
- Doetjes, Jenny, and Martin Honcoop. 1997. "The Semantics of Event-related Readings: a case for pair quantification." In *Ways of scope taking*, ed. by Anna Szabolcsi, 263–310. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5814-5_8
- Dotlacil, Jakub. 2015. "Why Is Distributivity so Hard? New Evidence from Distributive Markers and Licensors in Czech." In *Slavic Languages in the Perspective of Formal Grammar: Proceedings of FDSL 10.5, Brno 2014*, ed. by Markéta Ziková, Pavel Caha, and Mojmir Docekal, 139–154. Frankfurt: Peter Lang.
- Farkas, Donka. 2015. "Dependent Indefinites Revisited." Presentation at the Workshop on (Co-) distributivity, Paris.

- Henderson, Robert. 2016. "Variation in Dependent Indefinites." Presentation at the Workshop on (Co-)distributivity, Paris.
- Krifka, Manfred. 1990. "Four Thousand Ships Passed Through the Lock: Object-induced Measure Functions on Events." *Linguistics and Philosophy* 13: 487–519. <https://doi.org/10.1007/BF00627291>
- Oh, Sei-Rang. 2001. "Distributivity in an Event Semantics." In *11th Semantics and Linguistic Theory Conference (SALT 11)*, ed. by Rachel Hastings, Brendan Jackson, and Zsófia Zvolenszky, 326–345. Ithaca, NY: Cornell University. <https://doi.org/10.3765/salt.v11i0.2844>
- Oh, Sei-Rang. 2006. "Plurality Markers across Languages." Ph.D. dissertation, University of Connecticut.
- Panaitescu, Mara. To appear. "Romanian Dependent Numerals as Ratios." *Proceedings of Going Romance 2015*.
- Szabolcsi, Anna. 2010. *Quantification. Research Surveys in Linguistics*. Cambridge: Cambridge University Press.
- Szabolcsi, Anna. 2011. "Each What? Individual-key and Event-key Distribution." Presentation at the Workshop "Quantification", Stuttgart.
- Vlach, Frank. 1993. "Temporal Adverbials, Tenses and the Perfect." *Linguistics and Philosophy* 16 (3): 231–283. <https://doi.org/10.1007/bf00985970>
- Zhang, Linmin. 2013. "A Rate Analysis of Binominal Each". Poster at the Mid-Atlantic Colloquium of Studies in Meaning 3.
- Zimmermann, Malte. 2002. "Boys buying two sausages each: On the syntax and semantics of distance-distributivity." Ph.D. dissertation, University of Amsterdam.

Stressed enclitics are not weak pronouns

A plea for allomorphy

Diego Pescarini

CNRS – UMR 7320

The paper compares a morphophonological and a syntactic approach to stress shift and enclisis/proclisis asymmetries. The former analysis envisages the co-occurrence of multiple allomorphs with the same featural content, which are distributed in accordance with phonological rules. The latter analysis, by contrast, is based on the hypothesis that alternations are due to the co-occurrence of items belonging to different functional classes. I argue that the syntactic analyses put forth so far are not superior to previous morphophonological accounts. I acknowledge that the synchronic distribution of allomorphs does not result from productive phonological rules. Nonetheless, this is not *per se* a valid argument in favor of syntactic, class-based accounts, which must be advanced in compliance with Occam's razor.

Keywords: clitics, prosody, syntax, allomorphy

1. Introduction

Many languages exhibit a double series of pronouns: strong pronouns have the same distribution of DPs, while the distribution of weak/clitic pronouns is much more constrained. This often correlates with further prosodic properties, e.g. absence of stress. Pronouns are therefore organized into *classes*, which are modelled in terms of *inner* syntax: strong elements are conceived as extended phrases, while *clitics* correspond (at least in the latter stage of their derivation) to a deprived structure, possibly to a single head exhibiting an affix-like behavior. The alternation between classes, although modelled in terms of syntactic constituency, is ultimately encoded in the lexicon: functional elements are stored in the lexicon as triplets formed by a *syntactic (sub)tree*, containing a *bundle of φ -features*, associated with a *phonological exponent*.

In particular, Cardinaletti and Starke (1999) argue that at least three classes of pronouns must be envisaged. Dèchaine and Wiltschko (2002) reach a similar conclusion, but, building on a different set of diagnostics (predicate/argument asymmetries, binding, obviation, switch reference, etc.), they come up with a classification that cuts across Cardinaletti and Starke's, meaning that the typology is, at best, far more complicated than previously thought.

Besides classificatory issues (How many classes? Are our tests reliable?), a class-based approach poses several theoretical problems. First, the idea that classes are modelled in terms of *inner* syntax is promising but, in the end, difficult to test. The crucial point is how to disentangle properties hinging on the internal structure of pronouns from phenomena attributable to external, clausal factors. Second, almost all studies on the phenomenon build on the intuition that the structure of clitic/weak elements is a subtree corresponding to the lower structure of strong elements. However, as discussed in Poletto (2006), nothing really supports this view. Rather, the fact that clitic pronouns have case morphology seems to indicate that clitic/weak pronouns do in fact spell out outer layers of the DP structure, which strong pronouns do not lexicalise anymore.

In conclusion, the classification criteria, when confronted with an ample dataset, are not always consistent (Pescarini, to appear), the taxonomy – if any – is very complicated, the hypotheses that classes result from the inner syntax and that clitic/weak elements are pruned subtrees are, at best, not (completely) grounded. Generally speaking, the overall theory seems far from falsifiable and, as a result, in the current literature, terms like 'weak' or 'strong' are seldom used following the restrictive, technical definition given in works such as Cardinaletti and Starke (1999).

In what follows, I argue for a more conservative approach, in which morphophonological properties of pronominal elements are accounted for without postulating any connection with syntactic properties. Although acknowledging that the mapping of syntactic classes into morphophonological structures is a promising avenue of research, this paper aims to argue that a traditional analysis featuring allomorphs is not inferior to syntactic, class-based accounts.

The paper is organised as follows: Section 2 deals with the status of clitic pronouns bearing stress; Sections 3, 4, and 5 deal with proclisis/enclisis alternations in Neapolitan, Catalan, and French, respectively. Section 6 concludes.

2. Stressed enclitics

Cardinaletti and Starke (1999, sect. 3.2.8) and Cardinaletti (2015a) claim that weak elements, unlike clitics, can bear their own stress and be disyllabic. Weak

elements, however, are not expected to attract their host's stress. Conversely, this often happens in the case of enclitics, see (1), in particular with combinations of two or more clitics. Furthermore, stress shift may correlate with a morphological asymmetry between proclitic and enclitic forms, as in some cases the latter end up resembling strong pronouns, cf. (2).

- (1) *Finir-lù* Viozene (Rohlf's 1966: 442)
 to.end=it
 'to end it'
- (2) a. *Il me le donne* French
 he to.me= it= gives
 'He gives it to me.'
- b. *Donne-le-moi!*
 Give=it=to.me
 'Give it to me!'

The following table, from Ordóñez and Repetti (2006: 168), summarizes the patterns in which clitics may receive or attract stress (data from Kenstowicz 1991, Bafle 1992, 1994, Peperkamp 1996, 1997, Monachesi 1996, Loporcaro 2000, a.o.). Similar phenomena are attested in other Romance languages, such as Balearic varieties and Romanian (Martin Maiden, p.c.).

Table 1. (Ordóñez and Repetti 2006: 168)

(1)	imper.	dat. enclitic	acc. enclitic	two enclitics
I: <i>Stress Stability</i> (Ital. Span. Cat)	nárra	nárra-mi	nárra-la	nárra-mi-la
II: <i>Generalized Penultimate Stress Shift</i> (some Lucanian varieties)	nárra	narrá-mi	narrá-la	narra-mí-la
III: <i>Two-Clitic Penultimate Stress Shift</i> (some S. Ital. varieties, such as the one spoken in Naples)	nárra	nárra-mi	nárra-la	narra-mí-la
IV: <i>Mixed Penultimate Stress Shift</i> (some S. Ital. varieties, such as the one of Calvello)	nárra	nárra-mi narrá-mi	nárra-la narrá-la	narra-mí-la
V: <i>Final Stress Shift</i> (some varieties of Sardinian, Gascon)	nárra	narra-mí	narra-lá	narra-mi-lá

A brief remark is in order concerning present-day Neapolitan (type III), which, according to Ledgeway (2009: 34–35), is more akin to southern dialects like Calvello (type IV), where stress shift with a single enclitic is mandatory when the host is proparoxytone, optional otherwise (more on this in fn. 3):

- (3) a. *fràveca* + *la* → *fravəcàlla*
 make.IMP it.F Neapolitan
 ‘make it’
- b. *àssə* + *mə* → *àssəmə / assàmmə*
 let.IMP me Neapolitan
 ‘let me’

Given the above data, two kinds of explanation can be put forth: one in which clitics affect the prosodic hierarchy, thus triggering stress-shift (Peperkamp 1996, 1997, Loporcaro 2000). Alternatively, one may argue that the enclitics in (1), (2), and Table 1 are in fact weak elements bearing their own stress (Laenzlinger 1993, 1994, Ordóñez and Repetti 2006, 2014, Repetti 2016, Cardinaletti 2015a, 2015b).

Under the former approach, Loporcaro (2000) argued convincingly that Romance clitic pronouns, either proclitic or enclitic, can be regarded as syllables sister to a Prosodic Word and dominated by a recursive Prosodic Word, cf. (4) (see Selkirk 1995 for alternative configurations). Stress shift therefore result from a postlexical reassignment of stress.

- (4) (clitic (host)_{PW})_{PW} affixal clitic

A phonological account like (4) has two main advantages. First, it can account straightforwardly for cases in which the stress is dislodged, but does not fall on the clitic. This means that the pronoun does not carry stress *per se*, but its presence triggers prosodic restructuring.¹ This is particularly true for cases in which stress shift is conditioned by the proparoxytone/paroxytone stress pattern of the inner word, as in (3).

Second, a phonological analysis accounts for the fact that clitic sequences are more readily stressed than single enclitics, since combinations of two or more clitics form a new foot, see (5):

- (5) ((host)_{PW} (clitic clitic)_{FT})_{PW}

Conversely, the hypothesis that stressed enclitics are weak pronouns opens the door to a possible account of Type V languages, which are particularly puzzling under a purely phonological account. In fact, in dialects spoken at the Ligurian/France border such as Viozene in (6) (but the same holds for Mallorcan Catalan; Torres-Tamarit and Pons-Moll 2015) the stress always falls on the rightmost clitic (Rohlf's 1966: 442), although in these dialects the stress does not necessarily fall on

1. Bafle (1992, 1994) pointed out that both the root and the penultimate clitic bear primary stress as the *o* in *pòrtə* is open (and open mid vowels in Neapolitan are allowed only in syllables with primary stress) and the inner enclitic is subject to metaphony, which is typical of tonic vowels.

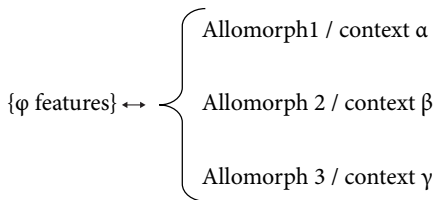
last syllable. This means that the above pattern cannot be easily derived from any language-specific constraint on stress assignment.

- (6) Finir-lù 'to end it'
 saver-lù 'to know it'
 portama-rù 'let us take it'
 vindirù 'sell it'
 servirsì 'to help oneself' Viozene, Ligurian

This and other facts led Ordóñez and Repetti (2006, 2014) to argue that stressed enclitics are in fact *weak pronouns* (see also Laenzlinger 1993, 1994 on French). The proposal is not straightforward as the above postverbal pronouns do not carry their own stress, but rather attract their hosts' stress and, to the best of my knowledge, no weak pronoun has never been reported to attract the stress of a nearby word. Furthermore, it is worth recalling that, under all other (syntactic) respects, e.g. doubling and resumption, the pronouns in (6) are in fact fully-fledged clitics.

However, as weak pronouns are more complex than clitics (Cardinaletti and Starke 1999: 178), the hypothesis that stressed enclitics are weak pronouns provides a promising explanation for some proclisis/enclisis asymmetries (more on this below). As weak pronouns are placed in a dedicated position of the clause that does not coincide with the one(s) of clitics (Cardinaletti 1991), Ordóñez and Repetti can account for the fact that stress shift always occur with enclitics, while proclitics are prosodically inert.

To summarize, this section has introduced two possible approaches to stress shift and enclisis/proclisis asymmetries. In the traditional morpho-phonological approach, alternations result from the co-occurrence of allomorphs, which derive from regular diachronic changes. In synchrony, allomorphs (1, 2, 3, etc.) are stored in the functional lexicon and eventually selected on the basis of the phonological context (α , β , γ , etc.):



Alternatively, one may argue that allomorphs do not exist as differences in shape always reflect differences in the inner syntax. Consequently, the distribution of exponents follows from the interplay between the inner and outer syntax of pronominal classes. Once introduced the two possible approaches, let us compare the two in the light of data from Neapolitan, French, and Catalan.

3. Neapolitan

In Neapolitan, proclitics are subject to processes of aphaeresis and elision, which reduce clitic clusters to a single syllable (see (7)), while enclitic clusters exhibit richer forms triggering stress shift. In particular, enclitic clusters have been argued to contain disyllabic reflexes of ILLE and INDE (Bafile 1992, 1994). Besides attracting stress, ILLE is subject to metaphony,² see (8): even if final vowels are reduced to -ə, the underlying ending triggers metaphony of the preceding vowel, which becomes -i- if the accusative clitic is masculine, and -e- if it is feminine.³

- (7) *t o 'pòrta* Neapolitan
 you= it= I.bring
 'I'll bring it to you.'
- (8) a. *pòrta=t=illə* Neapolitan
 bring=to.yourself=him/them.M/it.M
 'bring him/it.m/them for you'
- b. *pòrta=t=ellə*
 bring=to.yourself=her/them.F /it.F
 'bring her/it.f/them.f for you'

The fact that reflexes of ILLE and INDE attract stress more readily than first or second person clitics is consistent with the etymology. However, why are disyllabic outcomes of Latin pronouns preserved in enclisis? Bafile (1994: 16) argues that disyllabic exponents such as [i'llə] are retained in the functional lexicon because, once combined with another clitic form, they give rise to a paroxytone sequence. As for the contrast between enclitic [lə] and proclitic [o], it originated from a generalised process of centralisation of final vowels, therefore the former element is found enclitically, i.e. when it occurs at the right edge of a prosodic domain. Hence, even if the prosodic processes that led to the allomorphy are not productive anymore, the synchronic distribution of allomorphs still reflects the original

2. This metaphonetic distinction between the masculine and feminine is no longer very robust amongst most speakers and the originally non-metaphonetic form is generalized in most instances. See Ledgeway (2009) for extensive discussion.

3. The data in (7) and (8) are just part of the puzzle. An anonymous reviewer points out that “[t]he phonological approach predicts that in languages like Neapolitan, a form such as /rámmilə/ with antepenultimate stress should be grammatical (but instead is ungrammatical), and a form such as /frávəka-lə/ with pre-antepenultimate stress should be ungrammatical (but instead is grammatical). (NB. Both /frávəka-lə/ and /fravəká-llə/ are attested [...]).” On the latter remark, see the extensive discussion in Ledgeway (2009: 34–35). Regarding */rámmilə/, Bafile (1994) argues that they result from a ‘paradigmatic extension of the stress pattern typical of sequences containing two clitics’ (the translation is mine).

conditions. The list of allomorphs and relevant contexts are illustrated in (9); the order, as customary, goes from the form with the most restrictive distribution (e.g. [ˈillə]) to the default form, in compliance with the so-called Elsewhere Principle (or Panini’s Principle).

$$(9) \quad [\text{Pers: 3; Num: SG; Gen: M}] \leftrightarrow \left\{ \begin{array}{l} [\text{'illə}] / \text{clitic } _ \text{PrW} \\ [\text{lə}] / _ \text{PrW} \\ [\text{o}] \end{array} \right.$$

Ordóñez and Repetti (2014) argue for an alternative analysis in which Neapolitan clitics fall into three classes: besides ‘morphologically complex’ (*lə*) and ‘morphologically simple’ clitics (*o*), Neapolitan exhibits stressed, disyllabic weak pronouns (*illə*):

- (10) weak: *illə*
 complex clitic: *lə*
 simple clitic: *o*

In Cardinaletti and Starke’s (1999) and Cardinaletti’s (2008) formulation, complex/simple clitics and weak pronouns differ with respect to their syntactic make up and, consequently, in terms of their featural content. According to Cardinaletti and Starke (1999), strong pronouns feature an outer layer C, which allows the pronoun to be coordinated, modified, contrasted, etc. Clitics differ from weak elements in lacking a further layer (namely, Σ), whose absence correlates with syntactic and morphophonological properties, e.g. doubling, prosodic deficiency, etc. One can extend the same approach to the distinction between complex and simple clitics, which lack the *l* formative “marking definiteness” (Ordóñez and Repetti 2014: 176). For the sake of clarity, let us assume the following representation of the four classes introduced so far (I am slightly departing here from Ordóñez and Repetti’s own analysis):

- (11) a. Strong
 [C [Σ [D [Φ]]]]
 b. Weak
 [Σ [D [Φ]]]
 c. Complex Clitic
 [D [Φ]]
 d. Simple Clitic
 [Φ]

Hence, a tentative representation of Neapolitan ‘clitics’ is as follows:

- (12) $[\Sigma [D [\Phi]]] \leftrightarrow \text{ill}\partial$
 $[D [\Phi]] \leftrightarrow \text{l}\partial$
 $[\Phi] \leftrightarrow \text{'o}$

The elements in (12) are eventually attracted by different kinds of probes located in the functional spine of the clause. Crucially, weak elements are attracted to functional projections that are located below the probes attracting clitics, yielding the order *imperative clitic weak*. In Ordóñez and Repetti’s (2014) analysis, T and C probe clitics, while v is a probe for weak pronouns. For imperatives, they assume v -to-C movement and cliticization in C. Furthermore, to account for Neapolitan they assume a language-specific restriction constraining the number of clitics occurring in C: if a clitic is incorporated to a moved-to-C verb, then the other pronoun must be expressed by a weak element in v (“When C is not a possible probe for clitics in a given language, then v becomes a possible probe and triggers attraction of weak pronouns.” Ordóñez and Repetti 2014: 190). The derivation is as follows:

- (13) a. $[_C \text{ra}=\text{mm} \quad [_v = \text{ill}\partial]]$
 b. $*[_C \text{ra}=\text{m}=\text{l}\partial \quad [_v \dots]]$
 Give=to.me=it

The constraint against the co-occurrence of two clitics sounds rather ad hoc. Moreover, given (13), one may wonder about the position of aspectual adverbs corresponding to English *already*, *just*, *often*, etc. These adverbs occur at the v /T boundary and, in Italian, some of them precede the weak pronoun *loro*, cf. (14) (Cardinaletti 1991). Under the analysis in (13a), one might expect certain aspectual adverbs such as *mai* (‘never’) to occur between the clitic and weak pronouns, but, to the best of my knowledge, the prediction is not borne out.

- (14) *Non ho dato mai loro un libro.*
 not I.have given ever to.them a book
 ‘I never gave them a book.’

To conclude, the synchronic distribution of Neapolitan forms follows from phonological rules that are not productive anymore. Nonetheless, this does not necessarily mean that the distribution of present-day allomorphs is conditioned by syntactic factors. In particular, the above data do not support the claim that allomorphs belong to different classes.

4. Catalan

To bring further support to a class-based analysis of allomorphs, Ordóñez and Repetti (2014) focus on Catalan masculine accusative clitics: *el* and *lo*. The former contains an epenthetic vowel and occurs in proclisis, while the latter features a masculine singular ending and occurs in enclisis.

- (15) a. *El/*lo vol comprar*
 CL wants to.buy
 b. *Vol comprar-lo/*el*
 wants to.buy-CL

Again, Ordóñez and Repetti (2014) claim that “this restriction is not due to a phonological restriction [...] in fact, this form [*el*] is required with infinitives ending in *-re*.”

- (16) *Podries veure'l/*lo*
 you.could see it

Then, they claim that the *el/lo* alternation depends on the probing features occurring with tensed and untensed verbs. With the former, clitics are attracted by a probe endowed with a definite feature, but with no gender/number specification (this is the reason why only the simple clitic *el* is attracted). Conversely, with infinitives “ending in *-re*”, the probe has a richer set of features attracting the complex form *lo* (but why should verbs ending in *-re* exhibit a peculiar syntactic behavior?).

Alternatively, I argue that the distribution of the above outcomes was phonologically regular as it ultimately follows from the so-called Gröber Law (after Gröber 1877). A similar alternation occurs in Alguerese Catalan (Loporcaro 1997) or occurred in several northern Italo-Romance vernaculars (Gröber 1877, Vanelli 1992, 1998: 169–214).

Medieval Romance vernaculars featured a rule of apocope, by which word-final *-o/-e* are dropped after a single sonorant.⁴ Apocope also targets the *-o* of the clitic element *lo* (< ILLUM), yielding the form *l* in preconsonantal position (while in prevocalic position the allomorph *l* resulted from elision). Later on, the apocopated *l* was resyllabified by means of a prosthetic vowel *e* – yielding *el* – in

4. It is worth noting that apocope cannot be considered a merely phonological process since in many cases it is sensitive to the morpho-syntactic nature of the final vowel: in early Italian, for instance, if *-e* is a feminine plural ending, it never undergoes apocope, while apocope can take place if *-e* derives from the thematic vowel of the Latin 3rd declension. As the rule can discriminate between different kinds of inflectional endings, it means that the rule of apocope, even in Early Italo-romance, has a morphophonological nature. More on this in Pescarini (2013).

proclisis and before consonants. Hence, the diachronic evolution is as follows: $*(il) lo > l(o) > el /_C$.

The $l(o)/el$ alternation was eventually morphologized once the conditions ruling apocope and prosthesis ceased to exist. However, the conditions on the distribution of the various allomorphs in present-day languages still reflect the conditions under which apocope was triggered or blocked. Hence, I am not claiming that (synchronic) phonology is sufficient to account for the distribution of allomorphs; rather, I am suggesting that the synchronic distribution of allomorphs results from phonological conditions even if, synchronically, they do not hold anymore.

With this in mind, let us address the case of sequences formed by the enclitic lo and an infinitive. In early Romance, both elements (namely, the clitic and the preceding verb) may be subject to apocope. Therefore, a sequence formed by an infinitive, e.g. *fare* 'to make' and lo 'it/him,' could display two possible patterns of apocope: apocope of the infinitive or apocope of the clitic, cf. (17).

The former pattern is more likely because apocope targets the verb first (e.g. *fare* → *far*) and then apocope cannot target the prosodic constituent (verb+clitic)_{PrW} otherwise the resulting output would be syllabically illicit, e.g. $*farl$. This in turn explains why the non-apocoped form lo in languages such as central Catalan occurs only in enclisis after apocoped verbal forms.

- (17) a. *far(e) lo*
 b. *fare l(o)*
 do.INF =it/him.ACC
 'to do it/him'

Hence, apocope can target the enclitic pronoun if and only if the infinitive has not been apocoped, as in (17b). In central Catalan, in fact, the clitic is obligatorily apocoped after infinitives "ending with *-re*":

- (18) *Podries veure'l/*lo*
 you.could see it

To conclude, the enclitic lo resists apocope after an apocoped infinitive, as in (17a), and undergoes apocope after non-apocoped infinitives. In proclisis (and the same holds for the definite article), apocope is mandatory and the l formative is eventually syllabified by means of a prosthetic vowel, e.g. *el*. Nowadays, the distribution of lo , *el*, l allomorphs follows the same conditions although apocope is not active anymore. Then, given a sound historical explanation, I cannot see how the synchronic analysis would be improved by postulating the co-occurrence of different classes of elements, corresponding to different layers of structures, and probed by different syntactic heads.

5. French

In many languages, enclitic pronouns tend to be ‘heavier’ than proclitics even in absence of stress shift phenomena (see also Renzi and Vanelli 1983 on subject clitics). Several asymmetries can be accounted for under trivial phonological accounts, but other alternations are more puzzling. For instance, in modern French first and second person enclitics are expressed by an exponent identical to that of strong forms:

- (19) a. *Il me le donne*
 he to.me it gives
 ‘He gives it to me’
 b. *Donne-le-moi!*
 give-it-to.me
 ‘Give it to me!’

Laenzlinger (1993), Kayne (2003), Cardinaletti (2008) argue that proclitics are monomorphemic, while enclitics such as *moi* have a bipartite structure (e.g. *m-oi*). However, phonologically speaking, both forms are regular outcomes of the same monomorphemic Latin form *ME* in stressed and unstressed position respectively.

Laenzlinger (1993) argues that the *me/moi* alternation cannot depend on the assignment of stress to the word-final syllable (Foulet 1924). In fact, the same alternation is observed in non-standard varieties displaying the opposite order of clitics, e.g., *donne=moi=le*. Moreover, while *me/te* cannot occur enclitically, the third person clitic *le* is free to follow imperatives, see (15). According to Laenzlinger, this means that no phonological constraint requires enclitics to bear stress.

- (20) *Invite-le/*me*
 invite him/*me

Laenzlinger (1993) put forth the hypothesis that enclitics like *moi* are in fact weak pronouns, while Ordóñez and Repetti (2014) argue that forms like *moi/toi* are complex clitics (hence, bimorphemic elements, but not XPs) and that the elements probing clitics in French imperatives cannot attract simple clitics.

As in the case of Neapolitan and Catalan, however, the nature of the ban remains rather obscure and neither syntactic classification is supported by independent evidence. However, unlike the above data from Neapolitan and Catalan, the French situation is far more complicated as, to the best of my knowledge, no historical explanation can account straightforwardly for the current distribution of enclitic allomorphs.

6. Conclusions

In the previous sections I have compared a traditional/morphophonological analysis with an alternative syntactic approach to stress shift and enclisis/proclisis asymmetries. The former analysis envisages the co-occurrence of multiple allomorphs (having the same featural content) distributed in accordance with phonological rules. The latter analysis, by contrast, is based on the hypothesis that alternations are due to the co-occurrence of items belonging to different classes (thus, having different featural contents), while the nature and location of probes (subject to crosslinguistic variation) rule the distribution of clitic/weak elements.

First of all, it is worth noting that the alternation between weak, complex clitics, and simple clitics never gives rise to any peculiar semantic/pragmatic reading. Hence, although we end up postulating a different array of interpretable features for each allomorph, no relevant semantic effect has been detected. Moreover, the nature and position of the probes giving rise to linguistic variation cannot be ascertained by means of independent evidence. In conclusion, the syntactic analyses put forth so far does not seem to be more elegant or solve more puzzles than previous phonological accounts as the number of variables at play is very high and few of them can be controlled in order to verify or falsify the hypothesis.

On the contrary, a more traditional model can cope with the data once we abandon the (trivial) idea that synchronically active phonological rules can account for all the puzzles. In fact, many processes have been eventually morphologized, yielding alternations between lexical formatives, e.g. *le/il/le*, *li/lui*, *me/moi*, *el/lo* etc. that are synchronically opaque.

This, however, is not *per se* an argument in favor of syntactic, class-based accounts, which must be advanced in compliance with Occam's razor. I therefore do not exclude that the distribution of allomorphic variants is subject to syntactic factors, but I suggest that syntactic analyses should account for puzzling distributions, not for patterns that, from a historical point of view, are regular.

References

- Bafle, Laura. 1992. "Fonologia prosodica e teoria metrica: accento, cliticizzazione e innalzamento vocalico in napoletano." Ph.D. dissertation, Università di Firenze.
- Bafle, Laura. 1994. "La riassegnazione postlessicale dell'accento del napoletano." *Quaderni del Dipartimento di Linguistica - Università di Firenze* 5: 1–23.
- Cardinaletti, Anna. 1991. "On pronoun dative movement. The Italian dative *loro*." *Probus* 3: 127–153. <https://doi.org/10.1515/prbs.1991.3.2.127>

- Cardinaletti, Anna. 2008. "On different types of clitic clusters." In *The Bantu – Romance Connection. A comparative investigation of verbal agreement, DPs, and information structure*, ed. by Cécile De Cat and Katherine Demuth, 41–82. Amsterdam: John Benjamins.
<https://doi.org/10.1075/la.131.06car>
- Cardinaletti, Anna. 2015a. "Cases of apparent enclisis on past participles in Romance varieties." *Isogloss 1* (2): 179–197. <https://doi.org/10.5565/rev/isogloss.16>
- Cardinaletti, Anna. 2015b. "Syntactic Effects of Cliticization." in *Syntax. Theories and Analyses, Volume I*, ed. by Tibor Kiss, and Artemis Alexiadou, 595–653. Berlin: Mouton de Gruyter.
- Cardinaletti, Anna, and Michal Starke. 1999. "The typology of structural deficiency: A case study of the three classes of pronouns." In *Clitics in the Languages of Europe*, ed. by Henk van Riemsdijk, 145–233. Berlin: Mouton de Gruyter.
- Déchaine, Rose-Marie, and Martina Wiltschko. 2002. "Decomposing pronouns." *Linguistic Inquiry* 33: 409–422. <https://doi.org/10.1162/002438902760168554>
- Foulet, Lucien. 1924. "L'accent tonique et l'ordre des mots: formes faibles du pronom personnel après le verbe." *Romania* 50: 54–93.
- Gröber, Gustav. 1877. "'Lo', 'li' – 'il', 'i' im Altitalienischen." *Zeitschrift für romanische Philologie* 1: 108–110.
- Kayne, Richard. 2003. "Person Morphemes and Reflexives in Italian, French and Related Languages." In *The Syntax of Italian Dialects*, ed. by Christina Tortora, 102–136. New York: Oxford University Press.
- Kenstowicz, Michael. 1991. "Enclitic Accent: Latin, Macedonian, Italian, Polish." In *Certamen Phonologicum II*, ed. by Pier Marco Bertinetto, and Michele Loporcaro, 173–85. Turin: Rosenberg & Sellier.
- Laenzlinger, Christopher. 1993. "A syntactic view of Romance pronominal sequences." *Probus* 5 (3): 242–270. <https://doi.org/10.1515/prbs.1993.5.3.241>
- Laenzlinger, Christopher. 1994. "Enclitic clustering: The case of French positive imperatives." *Rivista di Grammatica Generativa* 19: 71–104.
- Ledgeway, Adam. 2009. *Grammatica diacronica del napoletano*. Tübingen: Niemeyer.
<https://doi.org/10.1515/9783484971288>
- Loporcaro, Michele. 1997. "On Vowel epenthesis in Alguer Catalan." In *Certamen Phonologicum III*, ed. by Pier Marco Bertinetto, Livio Gaeta, Georgi Jetchev, and David Michaels, 205–227. Torino: Rosenberg & Sellier.
- Loporcaro, Michele. 2000. "Stress stability under cliticization and the prosodic status of Romance clitics." In *Phonological Theory and the Dialects of Italy*, ed. by Lori Repetti, 137–168. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.212.09lop>
- Monachesi, Paola. 1996. "On the representation of Italian clitics." In *Interfaces in Phonology*, ed. by Ursula Kleinhenz, 83–101. Berlin: Akademie Verlag.
- Ordóñez, Francisco, and Lori Repetti. 2006. "Stressed Enclitics?" In *New Analyses on Romance Linguistics: Volume II: Phonetics, Phonology and Dialectology (Selected Papers from the 35th LSRL 35)*, ed. by Jean-Pierre Montreuil, 167–181. Amsterdam: John Benjamins.
- Ordóñez, Francisco, and Lori Repetti. 2014. "On the morphological restriction of hosting clitics in Italian and Sardinian dialects." *Italia dialettale* 75: 173–199.
- Peperkamp, Sharon. 1996. "On the Prosodic Representation of Clitics." In *Interfaces in Phonology*, ed. by Ursula Kleinhenz, 102–127. Berlin: Akademie Verlag.
- Peperkamp, Sharon. 1997. *Prosodic Words*. The Hague: Holland Academic Graphics.

- Pescarini, Diego. 2013. "Clitic clusters in early Italo-Romance and the syntax/phonology interface." In *Challenging Clitics*, ed. by Christine Meklenborg Salvesen, and Hans Petter Helland, 255–282 Amsterdam: John Benjamins. <https://doi.org/10.1075/la.206.11pes>
- Pescarini, Diego. To appear. "An emergentist view on functional classes." In *Linguistic Variation: Structure and Interpretation – A Festschrift in Honour of M. Rita Manzini in occasion of her 60th birthday*, ed. by Ludovico Franco, and Paolo Lorusso. Berlin: Mouton De Gruyter.
- Poletto, Cecilia. 2006. "Doubling as economy." *University of Venice Working Papers in Linguistics* 16: 211–235.
- Renzi, Lorenzo, and Laura Vanelli. 1983. "I pronomi soggetto in alcune varietà romanze." In *Scritti linguistici in onore di G.B. Pellegrini*, 120–145. Pisa: Pacini.
- Repetti, Lori. 2016. "The phonology of postverbal pronouns in Romance languages." In *Romance Linguistics 2013: Selected papers from the 43rd Linguistic Symposium on Romance Languages (LSRL)*, ed. by Christina Tortora, Marcel den Dikken, Ignacio L. Montoya, and Teresa O'Neill, 361–378 Amsterdam: John Benjamins. <https://doi.org/10.1075/rllt.9.19rep>
- Rohlf, Gerhard. 1966. *Grammatica storica della lingua italiana e dei suoi dialetti*. Turin: Einaudi.
- Selkirk, Elisabeth. 1995. "The prosodic structure of function words." In *Papers in Optimality Theory*, ed. by Jill Beckman, Laura Walsh Dickey, and Suzanne Urbanczyk, 439–70. Amherst, MA: GLSA Publications. [also in *Signal to Syntax: Bootstrapping from Speech to Grammar in Early Acquisition*, ed. by James L. Morgan, and Katherine Demuth, 187–214. Mahwah NJ: Lawrence Erlbaum Associates, 1996].
- Torres-Tamarit, Francesc, and Clàudia Pons-Moll. 2015. "Enclitic-triggered stress shift in Catalan." *Talk given at Going Romance 29*, University of Nijmegen, 10 December 2015.
- Vanelli, Laura. 1992. "Da lo a il: Storia dell'articolo definito maschile singolare in italiano e nei dialetti italiani." *Rivista italiana di dialettologia* 16: 29–66.
- Vanelli, Laura. 1998. *I dialetti italiani settentrionali nel panorama romanzo*. Rome: Bulzoni.

Causativization of verbs of directed motion in Romance languages

Anna Pineda
CNRS-IKER

The purpose of this paper is to shed light on a particular behavior displayed by verbs of inherently directed motion, the availability of a causative transitive alternate, and how this phenomenon interacts with the presence of clitic *se* in the intransitive variant. Data from Catalan, Aragonese and Italian varieties, none of which have received much attention in the literature, will prove crucial for our proposal. We adopt an inter-Romance perspective and a nanosyntactic approach to lexicalization in order to refine the correlation Jiménez-Fernández and Tubino (2014, 2017) find for Spanish, where inherently directional motion verbs are claimed to be more easily causativized (*entrar el coche* ‘go in the car’) in varieties where the use of *se* in the intransitive forms is also more frequent (*Juan se entró* ‘Juan *SE* went in’).

Keywords: syntactic microvariation, causativization, verbs of directed motion, clitic *se*, locative clitic, Catalan, Aragonese, Italian, Spanish.

1. Introduction

Two main types of unaccusative verbs have been distinguished in Romance languages. The first type corresponds mainly to change of state verbs entering the so-called (and extensively studied) *causative alternation*, exemplified in (1)–(2) for Catalan. Very often, the intransitive variant (1) bears the clitic *se*, which can be seen as an instance of anticausative morphology:¹

- (1) *La copa es trencà*
the glass *SE* broke.3SG
‘The glass *SE* broke.’

1. See Schäffer (2008) for an overview of contrasts across and within languages regarding the marking of (anti-)causatives.

- (2) *El convidat trencà la copa*
 the guest broke.3SG the glass
 ‘The guest broke the glass.’

The second type of unaccusative verbs, which we will focus on, does not normally bear *se* and has no causative counterpart:

- (3) *El Guiu vindrà/ arribarà/ naixerà/ entrarà/ pujarà/ baixarà*
 the Guiu will.{come/ arrive/ be.born/ come/ go.up/ go.down}.3SG
 ‘Guiu will come / arrive / be born / come in / go up / go down.’

Directed motion (from now on, DM) verbs are an important set of this second type of unaccusatives, and our goal is to explain an “unexpected” behavior they show in several Romance languages, namely their ability to have a causative transitive alternate.

2. Causative DM verbs

Transitive uses of DM verbs have not received much attention in Romance linguistics, probably because it is especially productive in non-standard varieties, as we shall see.²

2.1 Causativized DM verbs Catalan

For Catalan, the transitive use of DM verbs (4)–(6), common to all varieties, has been noted in some descriptive grammars (Moll 1990: 133–134, Badia i Margarit 1994: 287). To the best of our knowledge, only Gràcia (1989: 86) approaches this phenomenon from a theoretical perspective, although it was not the focus of her research.

- (4) *Martí, puja un parell de priorats del celler i*
 Martí, go.up.IMPER.2SG a couple of Priorats from.the wine cellar and
aprofita per baixar-hi aquestes caixes
 use.IMPER.2SG to go.down-LOC these boxes
 ‘Martí, bring a couple of Priorats from the wine cellar and take the opportunity to take those boxes down there’

2. For reasons of space we will not discuss French, where causativization of DM verbs is quite generalized (Rohlf 1954b: 10, Blinkenberg 1960: 118–122).

- (5) *Quan entraràs el cotxe al garatge?*
When come.in.FUT.2SG the car into.the garage
'When will you pull the car into the garage?'
- (6) *L'estudiant ha tornat el llibre a la biblioteca*
The student has gone.back the book to the library
'The student has returned the book to the library.'

2.2 Causativized DM verbs in Italian

Causativization of DM verbs is also possible in Italian:

- (7) *L'ho arrivato ('raggiunto')*
cl.ACC.3SG.MASC have.1SG arrived
'I have reached him.'
- (8) *Mi ha ritornato il libro*
CL.DAT.1SG has come.back the book
'(S)he has given the book back to me.' (Rohlf's 1954b: 10)

Rohlf's (1954b: 11) specifies that this phenomenon is much more productive in the dialects, especially in the south of Italy.³ Ledgeway (2000: 301, fn. 23) mentions that a number of unaccusative verbs "have developed a transitive pendant" in the dialects of southern Italy, and gives examples with 'go back', 'arrive', 'go up', 'go down' and other non-motion unaccusatives such as 'remain (something)' meaning 'leave (something)'. Here we give Andriani's (2011) examples from the variety of Bari:

- (9) *Laurə trasə la magħənə jinde o garagə / jəssə la magħənə dō*
Laurə goes.in the car into the garage / goes.out the car from
garəgə
the garage
'Laurə drives the car into the garage / drives the car out of the garage.'
- (10) *Luiggə salə / afənnə la bəfəclettə*
Luiggə goes.up/ goes.down the bicycle
'Luiggə brings up/brings down the bicycle' (Andriani 2011: 56, 67–68)

3. Sicilian and Calabrese (*Trásiri una vacca* 'Make a cow go in', *Nèsiri i vacchi* 'Make the cows go out'), Calabrese (*Scindilu* 'make him go down', *Nchiánalu* 'make him go up'), Abruzzese (*Ndrá* [something] 'Make [something] go in'), Pugliese (*Assi li vachhi* 'Make the cows go out'), Neapolitan (*Arrevà* [someone] 'Catch someone up'), among others (Rohlf's 1954b: 11).

2.3 Causativized DM verbs in Aragonese

In Aragonese causativized DM verbs are also found:⁴

- (11) *Chuan puyó una botella d'a bodega / baixó una botella*
 Chuan went.up.3SG a bottle from the cellar / went.down.3SG a bottle
t'a bodega
 to the cellar
 'Chuan carried up a bottle from the cellar / carried down a bottle to the cellar'
- (12) *Chuan entró o coche a la cochera*
 Chuan went.in.3SG the car into the garage
 'Chuan pulled the car to the garage'

2.4 Causativized DM verbs in Spanish

In Standard Spanish causative uses of DM verbs are not unknown (13), but the phenomenon is more pervasive (and condemned as colloquial⁵) in dialects of Andalucía and in Aragon, Avila and other northern Castile regions, as well as in certain regions of South America (14), where even other unaccusatives such as *caer* 'fall' and *quedar* 'remain' are causativized (Llorente Maldonado 1980: 31–32, Jiménez-Fernández and Tubino 2017).⁶

- (13) *Juan subió la botella de la bodega / bajó la botella a la bodega*
 Juan went.up.3SG the bottle from the cellar / went.down.3SG the bottle to the cellar
 'Juan carried up the bottle from the cellar / carried down the box to the basement'
- (14) *Sara entró el coche en el garaje*
 Sara went.in.3SG the car into the garage
 'Sara pulled the car into the garage'

4. Data provided by our informant Chorchi Díaz, member of the Academia de l'Aragonés.

5. For example, Llorente Maldonado (1980: 32) claims that they are 'colloquial structures that are absolutely illegitimate because *entrar* is a verb of movement and by its nature it is necessarily intransitive and it cannot be [sic] complemented by a direct object', but he admits that in some regions they are 'very common and no one born here, myself included, could ever exclude them from their spontaneous speech' [our translation].

6. See fn. 7.

3. What does a verb need to be causativized?

Jiménez-Fernández and Tubino (2014, 2017) note that in the Spanish spoken in Andalucía DM verbs are easily causativized. They also note that in this variety DM verbs appear with a clitic *se* more frequently than in Standard Spanish. Whereas in Standard Spanish *se* may appear with verbs such as *ir* ‘go’, *salir* ‘go out’, *venir* ‘come’, *bajar* ‘go down’ and *subir* ‘go up’, in Andalusian Spanish a wider range of DM verbs can take the clitic, including *entrar* ‘go in’, *llegar* ‘arrive’ and *huir* ‘go away’. In fact, a similar situation is found in other peninsular varieties, as well as in American Spanish (Martín Zorraquino 1979: 280–281).

Focusing on Andalusian Spanish, where most DM verbs can take *se* and can easily be causativized, Jiménez-Fernández and Tubino (2014, 2017) conclude that a correlation exists, as we summarize here in Table 1 for Standard Spanish and Table 2 for Andalusian Spanish.⁷

Table 1. Standard Spanish

Standard Spanish	<i>ir</i>	<i>caer</i>	<i>subir</i>	<i>bajar</i>	<i>entrar</i>
<i>se</i> -intransitive	✓	✓	✓	✓	✗
causative transitive	✗	✗	✓	✓	✗

Table 2. Andalusian Spanish

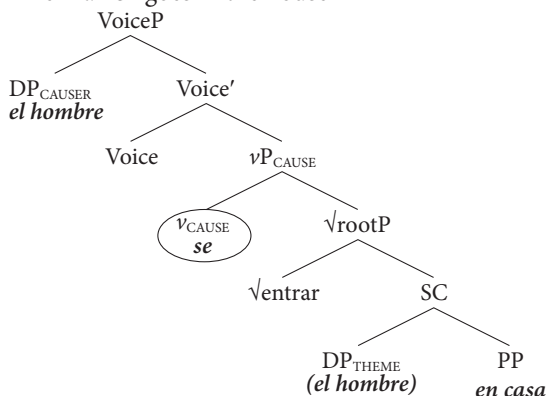
Andalusian Spanish	<i>ir</i>	<i>caer</i>	<i>subir</i>	<i>bajar</i>	<i>entrar</i>
<i>se</i> -intransitive	✓	✓	✓	✓	✓
causative transitive	✗	✓	✓	✓	✓

The tables above describe the following generalization: if a DM has a causative variant, the existence of a pronominal (i.e. formally reflexive) intransitive variant is “expected”. Let us see how the causative and the pronominal uses of these verbs are connected. According to De Miguel and Fernández Lagunilla (2000), the presence of *se* signals a particular event structure: whereas non-pronominal uses of DM verbs name only the culmination point or denote just the achievement

7. Martín Zorraquino (1979: 291–293) already claimed that [our translation] ‘fluctuations in the use of reflexives are not an isolated phenomenon, but must be studied together with changes and modifications in argument structure that occur with other Spanish verbs in both America and Spain’, giving examples of the transitive use of intransitive verbs such as *entrar*, *caer* (*has caído el paraguas* ‘(lit.) you have fallen the umbrella’), *quedar* (*te quedas todo en casa* ‘(lit.) you remain everything at home’), *correr* (*correr la bici* ‘run the bicycle’), a.o. See also Gràcia (1989: 90, fn. 12).

of a particular position, when combined with *se* they indicate a complex event structure formed by a causing subevent and a resulting subevent, as shown in (15).⁸

- (15) a. *El hombre se entra en casa*
 ‘The man *SE* goes in the house’
 b.



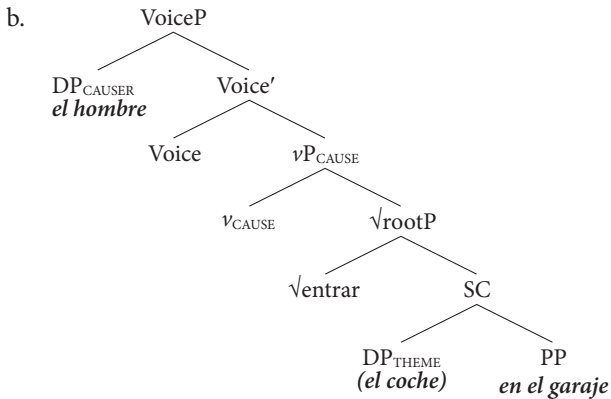
(Adapted from Jiménez-Fernández and Tubino 2017)

In this configuration, *se* is the morphosyntactic expression of the dual role of the DP subject (undergoer of the change of location and holder of the new state of being located in some place). In other words, *se* “is the realization of v_{CAUSE} as an indication that, even though the same referent [*el hombre*] is associated with both resulting and causing subevents, this must be interpreted as two separate entities,” and this “is the only way the structure may entail a change of state/location rather than a pure achievement” (Jiménez-Fernández and Tubino 2017).

The crucial point is that the bieventive structure in (15) paves the way for the causative transitive one. This follows from observations such as Schäfer’s (2012), according to whom what triggers causation is the presence of a resulting state in the inner composition of verbs or the syntactic compatibility with it of the resulting state. Thus, the complex structure in (15) with a resulting subevent gives rise to the causative structure (16), where two different entities (*el hombre* and *el coche*) are now involved in each subevent:

- (16) a. *El hombre entra el coche en el garaje*
 ‘The man pulls the car into the garage’

8. Following Folli and Harley (2005), in Jiménez-Fernández and Tubino’s (2014, 2017) approach, it is assumed that Causers are not necessarily [-human] (such as natural forces), but can also be [+human].



(Adapted from Jiménez-Fernández and Tubino 2017)

Essentially, the only difference between (15) and (16) is that in the former the same participant is related to the two subevents, whereas in the latter each subevent has its own participant.

What lies behind Jiménez-Fernández and Tubino's analysis is that causative transitive uses of DM verbs follow straightforwardly from the compatibility of roots denoting directed motion with a causative/agentive verbalizing head that may take a resulting state (see also Schäfer 2008, 2012). As a good example of this, consider the case of Spanish *entrar* and *llegar*: according to De Miguel and Fernández Lagunilla (2000: 30), these verbs should not admit *se* because [our translation] 'they simply focalize the point in which they occur'; however, given that they do so in Andalusian Spanish, Jiménez-Fernández and Tubino (2014, 2017) conclude that in this variety the conceptualization of these roots is different: *entrarse* would mean 'go in somewhere and remain in this new location' (i.e. with a resulting state/location) (15), and similarly for *llegarse*. In turn, this perceived compatibility of particular roots with the expression of a resulting state is the key element that allows causativization (16) (e.g. Spanish *entrar el coche*).

Adopting a more comprehensive Romance perspective, our purpose is to extend the connection between causativization of a verb and its compatibility with *se* to Catalan, Italian and Aragonese varieties which, as shown in Section 2, allow DM verbs to be causativized. In our analysis, what lies behind cross- and intra-linguistic differences such as the ones pointed out by Jiménez-Fernández and Tubino is the possibility for roots to lexicalize different portions of structure, as will be shown in Section 4.

We thus assume that speakers of Romance varieties where DM verbs can be causativized can make use of those roots not only to lexicalize a structure denoting the punctual achievement of a particular position (intransitive without *se*), but also a structure entailing a complex change of location event with a resulting

state (denoting permanence in a position).⁹ Therefore, our assumption is that all Romance varieties where causativization of DM verbs is possible have the bieventive structure in (15) that paves the way for the causative transitive structure in (16).

However, when we adopt a broader perspective we find that Jiménez-Fernández and Tubino's observations for Spanish may not be clearly generalizable in a cross-Romance perspective. In Catalan, it seems that the correlation seen in Tables 1 and 2 fails to hold, since *pujar*, *baixar* and *entrar* are normally used as causative verbs in all varieties (recall Section 2.1), but there is no pronominal intransitive form generally available (see shaded cells):

Table 3. Catalan

Catalan	<i>anar</i>	<i>caure</i>	<i>pujar</i>	<i>baixar</i>	<i>entrar</i>	<i>tornar</i>
se-intransitive	✓	x	x	x	x	✓
causative transitive	x	x	✓	✓	✓	✓

Contrary to what Catalan data seem to convey at first sight, we will argue that the generalization does not fail but needs to be refined. After looking at the distribution of pronominal DM verbs across Catalan, Aragonese and Italian varieties in more detail (Section 3), we will claim that the bieventive structure in (15), which we shall modify to accommodate our cross-linguistic data, is available for all the Romance varieties studied here even though the clitic is not always overt (Section 4). Only under this assumption can we account for the availability of causativization in all the varieties studied. We will have to explain, then, why Spanish distinguishes between *subir* (punctual, monoeventive) and *subirse* (complex, bieventive, as in (15)) whereas in Catalan, according to our proposal, the monoeventive and bieventive structures both correspond to the same form, *pujar* (or, more specifically, *pujar* for the monoeventive and *pujar*Ø for the bieventive). In our view, this cross-linguistic contrast is due to differences in the lexicalization pattern of the bieventive structure.

3.1 Pronominal DM verbs in (varieties of) Catalan

Most descriptive grammars of Catalan note that *anar* 'go' and *tornar* 'go back' can appear with the clitic cluster *se'n*, composed of the reflexive or aspectual clitic *se*

9. We do not talk about "verbs" but about roots and verbalizing heads instead, since we do not consider that verbs are listed in the lexicon with all their features specified, but instead, in the spirit of the *Distributed Morphology* neoconstructionist approach, we assume that verbs are formed in the syntax by combination of a root and a verbalizing head *v*.

and the locative ablative clitic *en* (Badia i Margarit 1994: 295). The forms *anar-se'n* and *tornar-se'n* are common to all varieties (Table 3).

In some Catalan dialects, however, virtually all DM verbs can take *se'n*: *venir-se'n* ('come'), *pujar-se'n* ('go up'), *baixar-se'n* ('go down'), *entrar-se'n* ('go in'), *eixir-se'n* ('go out'), among others. This is the case of Valencian Catalan (Sanchis Guarner 1950; Todolí 1990, 2002; Solà 1994) as well as the variety spoken in La Franja in Aragon (Giralt 1995). In the rest of the linguistic domain, these forms are only found in folk storytelling, proverbs and religious prayers.

For most Catalan speakers, then, the pronominal use of DM verbs is seen as an archaism. In Old Catalan the use of *se'n* was indeed possible with virtually all DM verbs:¹⁰

- (17) *dits senyors concellers [...]*
se·n, baixaren,
 REFL.3PL ABL went.down.3SG
a peu, baix a la vora de la mar (16th c.)
 'The aforementioned lords descended to the coast on foot'
- (18) *E, quan haguem sopat,*
pujam-nos-en
 went.up.1PL-REFL.1PL-ABL
en lo terrat del castell de Cuylera (14th c.)
 'And when we had had dinner, we went up to the terrace of Cuylera Castle'
- (19) *dos grans barons [...]*
vengueren-se-n
 came.3PL-REFL.3PL-ABL
là ún lo traydor era (14th c.)
 'Two great barons [...] came to where the traitor was'
- (20) *les gèns e ·ls cavalers [...]*
se·n entraren
 REFL.3PL ABL went.in.3PL
en la ciutat (13th c.)
 'the people and the knights [...] entered the city'

3.2 Pronominal DM verbs in (varieties of) Italian

For Italian, Rohlfs (1954a: 188) claims that in 'the literary language' the clitic *si* may appear with DM verbs such as *andarsi* ('go, leave'), *venirsi* ('come') or *uscirsi*

10. Examples are from the *Corpus Informatizat del Català Antic*, www.cica.cat.

(‘go out’). Interestingly, the translator of Rohlfs’ work (S. Persichino) specifies that the language prefers the pronominal construction (or rather admits it alone) when the reflexive clitic appears together with *ne* [our translation]: ‘No one would say, reflexively, *ci andiamo* [...] whereas [...], *ce n’andiamo* [...] *me ne venivo* and so on are in common use’ (Rohlfs 1954a: 188, fn. 1). Elsewhere, Rohlfs (1954b: 12, 1954a: 166) himself cites forms with the whole *se+ne* clitic cluster.

As in Catalan, in the case of Italian there are also some dialectal differences concerning the extension of the use of pronominal DM verbs. Southern dialects are especially flexible on this point. Luigi Andriani (p.c.) confirms this for Barese and regional Italian from Bari. Likewise, Neapolitan exhibits a variety of pronominal DM verbs:

- (21) *Me ne saglio a [...]*
REFL.1SG ABL go.up to
‘I go up to...’
- (22) *Po’ me ne tornaie ccà*
REFL.1SG ABL came.back.1SG
‘Then I came back here’
- (23) *Tu te ne viene, Federi??*
REFL.2SG ABL come.2SG
‘Are you coming (with us), Federi?’ (Ledgeway 2009: 313, 350)

3.3 Pronominal DM verbs in (varieties of) Aragonese

In Aragonese, the use of DM verbs with the clitic *se* and an ablative locative clitic has been common since an early stage of the language and is present, in almost the whole domain, for ‘go’ and ‘go down’, for example (Alvar 1953: 298). Again, a more extended range of verbs is found once we turn to varieties. In Eastern Aragonese, we find examples with several verbs:

- (24) *Ha dicho tío que te’n venises*
REFL.2SG ABL come.SUBJ.PAST.2SG
‘The uncle has said that you should come.’
- (25) *Súbitene t’ astirriba*
go.up.IMPER.2SG+REFL.2SG+ABL
‘Go up to the top.’
- (26) *Bájatene*
go.down.IMPER.2SG+REFL.2SG+ABL
‘Come down.’

- (27) *Salítone*
 go.out.IMPER.2SG+REFL.2SG+ABL
 ‘Go out.’
- (28) *La mullé ya se’n ha torna uta casa*
 The woman already REFL.3SG ABL has gone.back to home
 ‘The woman has already gone back home.’ (Arnal Purroy 1998: 316–317)

In Central Aragonese, this pattern is very frequent too:

- (29) *Cuando quieras í-te ne, t’én bas*
 go-REFL.2SG ABL REFL.2SG ABL go.PRES.2SG
 ‘When you want to go/leave, you go/leave.’
- (30) *Baxa te ne*
 come.IMPER.2SG REFL.2SG ABL
 ‘Come down.’
- (31) *M’én puyo ta ras cumbres más altas*
 REFL.1SG ABL go.up.1SG
 ‘I climb the highest peaks.’
- (32) *Nos ne tornaremos*
 REFL.1PL ABL will.go.back.1PL
 ‘We will come back.’ (Nagore Lain 1986: 108–109)

Finally, for Western Aragonese, examples with the clitic cluster (33) are found together with examples with only *se* (34):

- (33) *Ellos sen son tornaus*
 REFL.3PL+ABL are gone.back
 ‘They have gone back.’
- (34) *Te yeras tornado*
 REFL.2SG were.2SG gone.back
 ‘You had gone back.’ (Alvar 1953: 292)

4. Proposal

4.1 What is NE?

The Catalan, Italian and Aragonese examples above differ from Spanish in that DM verbs take not only the “reflexive” clitic *se* but also the ablative locative clitic *ne*. Descriptive grammars agree that in this context this element has lost its

pronominal function because it does not replace anything and just co-occurs with the verb. Nonetheless, in these varieties a full-fledged, truly pronominal ablative *ne* exists and can replace a source complement (i.e. ‘Does Anna come from Paris? Yes, she *NE* comes’).

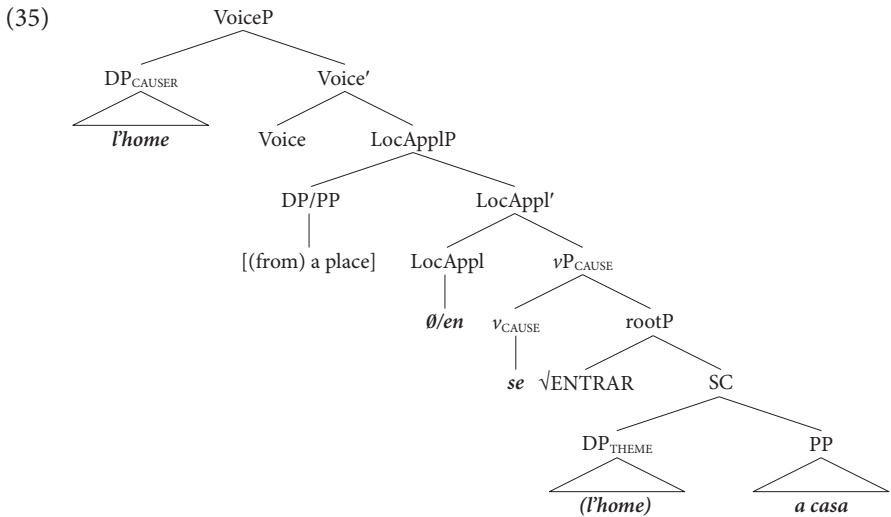
To find out what *ne* contributes to pronominal DM verbs, first it is crucial to see what has been said for Spanish, where only *se* appears with these verbs (compare Sp. *irse* vs. Cat. *anar-se’n*). We will argue that some claims based only on Spanish data need to be refined to accommodate data from other Romance languages: in particular, we will show that some properties that have been attributed to the presence of *se* should be attributed instead to *ne*.

As is well known, in Spanish an assortment of (transitive and intransitive) verbs can take the so-called aspectual *se* (e.g. *comerse un pastel* ‘eat.SE a cake’, *caerse* ‘fall.SE’). The term *aspectual* reflects the idea that *se* contributes an aspectual value, in the sense that the clitic can only occur with verbs denoting “bounded” events (Nishida 1994, Zagona 1996, De Miguel 1999, De Miguel and Fernández Lagunilla 2000, Basilico 2010, a.o.). Regarding the particular case of DM verbs, aside from referring the reader to Pineda (to appear), here we can only observe that most authors relate the presence of *se* with these verbs to the idea of a place that is left behind (therefore, events are bounded by this delimitation of the initial location). De Miguel (1999: 2995) claims that [our translation] ‘the interpretation is that there is an implicit or explicit mention of the beginning of the event, a mention that delimits the event’; Masullo (1992: 246) argues that *se* stands for an incorporated source argument; a similar idea is defended by Schrotten (1972: 89), according to whom “the most plausible suggestion is that the reflexive ‘variants’ imply movement from some place (that need not be specified explicitly), whereas the non-reflexive ‘variants’ do not imply such movement from some place”; along the same lines, Sánchez López (2002: 118) argues that DM verbs with *se* add a meaning of leaving behind the place of departure.

To sum up, it has been said that *se* with DM verbs in Spanish relates to the boundedness of the event, generally in the point of departure (for further details, see Pineda to appear). However, we think that an account of the role of *se* based only on Spanish data does not suffice and needs to be modified to cover the data in other Romance languages: in order to determine the contribution of *se* in Romance DM verbs, we must take into account the ablative locative clitic that appears, together with *se*, in Catalan, Italian and Aragonese varieties. The crucial point is the following: the fact that in these languages *se* appears together with an ablative/source clitic leads us to reconsider the idea that *se* stands for an incorporated source argument or references the source of the movement described by the event, as claimed by the authors mentioned above. Rather, adopting an inter-Romance perspective, we propose that it is the locative clitic that instantiates a syntactic

head whose contribution has to do with the notion of leaving behind a location. This fits with the ablative nature of the element, and also with the idea, noted in several descriptive grammars (and dictionaries), that DM verbs with *se+ne* have a meaning of ‘going away’ that the equivalent non-pronominal verbs do not have. It seems reasonable and intuitive to attribute this semantic notion to the presence of a locative source clitic *en*, rather than to *se*, as authors working only on Spanish data had previously concluded.

Following Pineda (to appear), we claim that this syntactic head spelled out as *ne* is a locative applicative:¹¹



Note that the locative head selects for the causative head, replicating the semantic relation between source and causation: causation is related to the presence of a resulting state, and at the same time this resulting state necessarily follows from leaving a place behind, thus the notion of source actually determines that there is a component of causation (broadly understood as the origin of a process of change).¹²

The locative (applicative) head can be phonologically overt (e.g. Catalan *en*) or covert (e.g. Spanish \emptyset), since we assume (within a *Distributed Morphology* approach) that the morphophonological realization of this head is regulated by

11. Owing to space constraints, we must leave all details aside and refer the reader to Pineda (to appear), where Bantu locative applicatives and those here proposed for Romance languages are compared.

12. I thank the two anonymous reviewers of this paper for having guided me towards this observation.

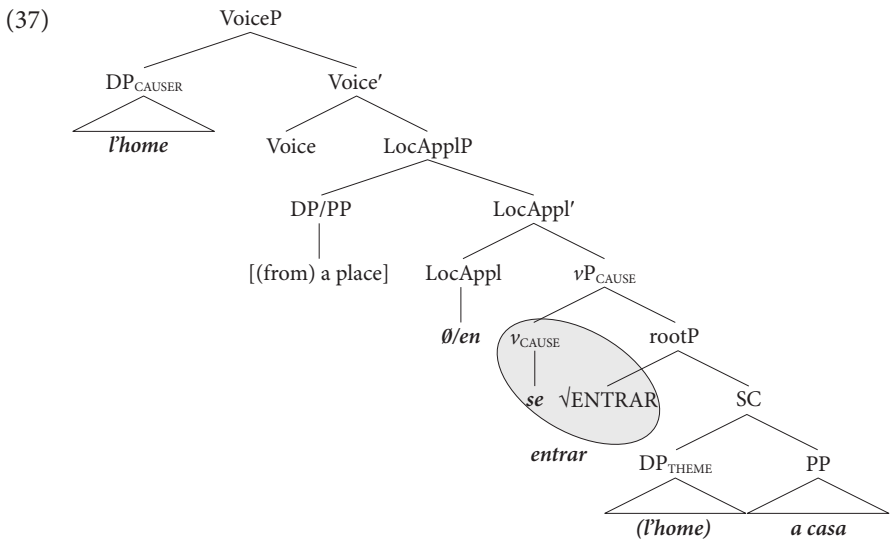
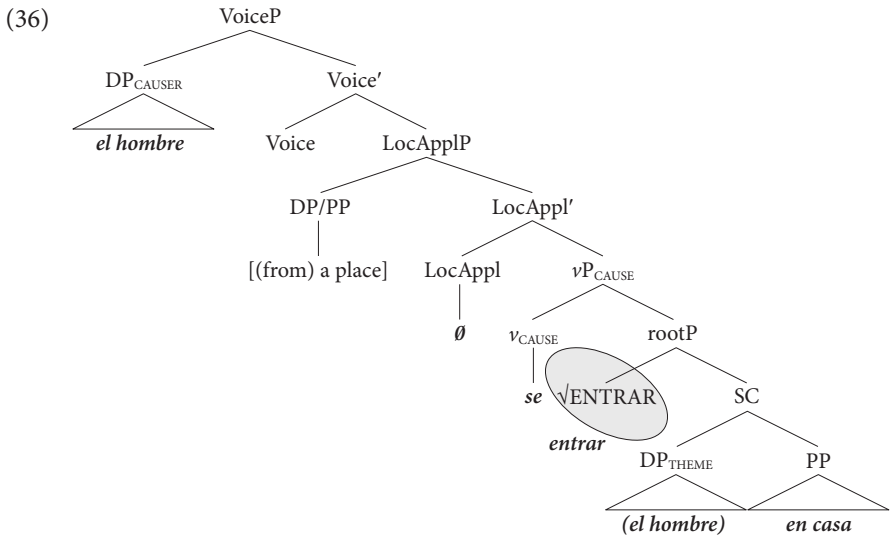
different rules across languages and even dialects. We propose that Spanish pronominal DM verbs do not have the structure in (15) but that in (35), with the locative syntactic head remaining covert. We thus refine Jiménez-Fernández and Tubino's proposal and offer a unified analysis that accommodates data not only from Spanish, but also from other Romance languages.

4.2 Same structure, different lexicalization

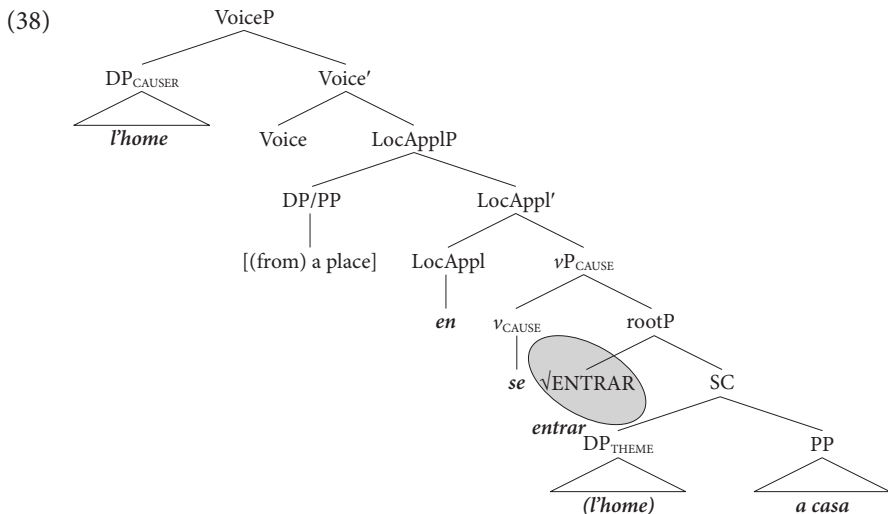
Variation is not limited to the realization of the locative head, though. Recall that our proposal is that all varieties with causative uses of DM verbs have the bieventive structure in (35), because this paves the way for subsequent causativization. However, this correlation seems not to be systematic. Let us take the case of *entrar*: in Andalusian Spanish, it can be causativized and we find the pronominal form *entrarse*, and in general Catalan it can be causativized too but we do not find a pronominal form, although we do find it in Valencian and La Franja Catalan (*entrar-se'n*). We thus have to explain the cross-linguistic and cross-dialectal contrasts regarding *se/Ø*.

We argue that this is due to different lexicalization patterns: whereas in Andalusian Spanish (36) each head (root and v_{CAUSE}) is lexicalized independently (*entrar* and *se*), in general Catalan (37) the same lexical item (*entrar*) lexicalizes both heads, instantiating the so-called Phrasal Spell Out, as proposed by the nano-syntactic approach to lexicalization.¹³

13. A similar solution we could resort to is Fusion, an operation invoked within the Distributed Morphology approach by virtue of which two nodes are fused into a single terminal node and then spelled out as one.



Finally, Valencian and La Franja Catalan behave like Andalusian Spanish, but with the locative head phonologically overt:



If we compare (36) for *entrarse* (and cross-linguistic equivalents) and (38) for *entrar-se'n* (and cross-linguistic equivalents), we observe that the overt realization of the locative head is subject to the presence of *se* expressly spelling out v_{CAUSE} . In other words, we cannot have *en* without *se*, but we can have *se* without *en*, as in Spanish or in some Aragonese examples as (34). Indeed, the diachronic evolution from Latin to Romance shows that the reflexive clitic with intransitive verbs was already possible in Latin, and more widespread in Vulgar Latin, whereas the locative clitic was grammaticalized later (see Pineda to appear and references therein).

Our analysis accounts for the substantial amount of variation across and within languages regarding DM verbs. The locative (applicative) head may be overt or covert, comparably to what has been proposed for dative applicative heads in Romance (Pineda 2013). On the other hand, the verbal root and the v_{CAUSE} may each have a dedicated exponent (e.g. *entrar* + *se*) or, alternatively, lexicalization of more than one head (Phrasal Spell Out in nanosyntactic terms) may take place, yielding one lexical item for both nodes (e.g. *entrar*).

5. Conclusions

In this work, we have analyzed an *a priori* unexpected behavior of DM verbs: their causative transitive use in several Romance varieties. We have adopted an inter-Romance perspective in order to contribute evidence showing that the ability of a particular root to lexicalize a causative structure is connected to the availability of a bieventive structure including a resulting state, which in turn also yields the pronominal uses of such verbs (even if we do not always see the pronominal form).

Crucially, our data from different Romance varieties has allowed us to refine the settings of this connection, arguing for the existence of different lexicalization patterns and positing a locative (applicative) head.

Acknowledgements

This work has been supported by the postdoctoral research fellowship Beatriu de Pinós 2014 BP_A 00165 (AGAUR, Generalitat de Catalunya) and the research project FFI2014-56968-C4-1-P (Ministerio de Economía y Competitividad, Spanish Government). I would like to thank the audiences of LSRL 46 (Stony Brook University, 2016) and CamComS 5 (University of Cambridge, 2016), for valuable comments on previous versions of this study. All errors remain mine.

References

- Alvar, Manuel. 1953. *El dialecto aragonés*. Madrid: Gredos.
- Andriani, Luigi. 2011. “Differential Object Marking, Clitic Doubling and Argument Structure in Barese.” M.A. thesis, Leiden University.
- Arnal Purroy, María Luisa. 1998. *El habla de la Baja Ribagorza occidental: aspectos fónicos y gramaticales*. Zaragoza: Institución “Fernando el Católico”.
- Badia i Margarit, Antoni Maria. 1994. *Gramàtica de la llengua catalana. Descriptiva, normativa, diatòpica i diastràtica*. Barcelona: Enciclopèdia Catalana.
- Basilico, David. 2010. “The *se* clitic and its relationship to paths.” *Probus* 22: 271–302. <https://doi.org/10.1515/prbs.2010.010>
- Blinkenberg, Andreas. 1960. *Le problème de la transitivité en français moderne*. Kobenhavn: Ejnar Munksgaard.
- De Miguel, Elena. 1999. “El aspecto léxico.” In *Gramática descriptiva de la lengua española*, ed. by Ignacio Bosque, and Violeta Demonte, vol. 2, 2977–3060. Madrid: Espasa-Calpe.
- De Miguel, Elena, and Marina Fernández Lagunilla. 2000. “El operador aspectual ‘se.’” *Revista española de lingüística* 30: 13–44.
- Folli, Raffaella, and Heidi Harley. 2005. “Flavours of *v*: Consuming Results in Italian and English.” In *Aspectual Inquiries*, ed. by Paula Kempchinsky, and Roumyana Slabakova, 95–120. Dordrecht: Springer. https://doi.org/10.1007/1-4020-3033-9_5
- Giralt, Javier. 1995. “Valors i funcions d’*IBIINDE* als parlars lliterans.” *Alazet* 7: 57–74.
- Gràcia, Lluïsa. 1989. *Els verbs ergatius en català*. Maó: Institut Menorquí d’Estudis.
- Jiménez-Fernández, Ángel, and Mercedes Tubino, 2014. “Variación sintáctica en la causativización léxica.” *RSEL* 44 (1): 7–37.
- Jiménez-Fernández, Ángel, and Mercedes Tubino. 2017. “Causativity in Southern Peninsular Spanish.” In *The Syntactic Variation of Spanish Dialects*, ed. by Ángel Gallego. Oxford: Oxford University Press.
- Ledgeway, Adam. 2000. *A Comparative Syntax of the Dialects of Southern Italy: A Minimalist Approach*. Oxford: Blackwell.
- Ledgeway, Adam. 2009. *Grammatica diacronica del napoletano*. Tübingen: Max Niemeyer Verlag. <https://doi.org/10.1515/9783484971288>

- Llorente Maldonado, Antonio. 1980. "Consideraciones sobre el español actual." *Anuario de Letras* XVIII: 5–61.
- Martín Zorraquino, M. Antonia. 1979. *Las construcciones pronominales en español*. Madrid: Gredos.
- Masullo, Pascual José. 1992. "Incorporation and case theory in Spanish. A cross-linguistic perspective." Ph.D. dissertation, University of Washington.
- Moll, Francesc de Borja. 1990 [1968]. *Gramàtica catalana referida especialment a les Illes Balears*. Palma de Mallorca: Moll.
- Nagore Lain, Francho. 1986. *El aragonés de Panticosa. Gramàtica*. Huesca: Diputació Provincial.
- Nishida, Chiyo. 1994. "The Spanish reflexive clitic *se* as an aspectual class marker." *Linguistics* 23: 425–458.
- Pineda, Anna. 2013. "Double Object Constructions in Spanish (and Catalan) Revisited." In *Romance Language and Linguistic Theory 2011*, ed. by Sergio Baauw, Frank Drijkoningen, Luisa Meroni, and Manuela Pinto, 193–216. Amsterdam: John Benjamins. <https://doi.org/10.1075/rllt.5.10pin>
- Pineda, Anna. To appear. "The role of SE and NE in Romance verbs of directed motion. Evidence from Catalan, Italian, Aragonese and Spanish varieties." In *Proceedings from the Workshop on Romance Se-Si*, ed. by Grant Armstrong, and Jonathan MacDonald.
- Rohlf, Gerhard. 1954a. *Grammatica storica della lingua italiana e dei suoi dialetti. Vol. 1. Morfologia*. Torino: Einaudi.
- Rohlf, Gerhard. 1954b. *Grammatica storica della lingua italiana e dei suoi dialetti. Vol. 2. Sintassi e formazione delle parole*. Torino: Einaudi.
- Sánchez López, Cristina. 2002. "Las construcciones con *se*. Estado de la cuestión." In *Las construcciones con se*, ed. by Cristina Sánchez López, 13–163. Madrid: Visor Libros.
- Sanchis Guarner, Manuel. 1950. *Gramàtica valenciana*. València: Torre.
- Schäfer, Florian. 2008. *The Syntax of Anti-Causatives. External Arguments in Change-of-state Contexts*. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.126>
- Schäfer, Florian. 2012. "Two Types of External Argument Licensing – The Case of Causers." *Studia Linguistica* 66 (2): 1–53. <https://doi.org/10.1111/j.1467-9582.2012.01192.x>
- Schroten, Jan. 1972. *Concerning the Deep Structure of Spanish Reflexive Sentences*. The Hague: Janua Linguarum.
- Solà, Joan. 1994. *Sintaxi normativa: estat de la qüestió*. Barcelona: Empúries.
- Todolí, Júlia. 1990. "Les funcions pronominals en el català del País Valencià." *Miscel·lània Joan Bastardas* 4, 253–261.
- Todolí, Júlia. 2002. "Els pronoms." In *Gramàtica del català contemporani*, ed. by Joan Solà, Maria-Rosa Lloret, Joan Mascaró, and Manuel Pérez Saldanya, vol. 2, 1337–1433. Barcelona: Empúries.
- Zagona, Karen. 1996. "Compositionality of Aspect: Evidence from Spanish Aspectual *se*." In *Aspects of Romance Linguistics*, ed. by Claudia Parodi, 475–488. Washington, DC: Georgetown University Press.

Latin denominal deponents

A syntactic analysis

Francesco Pinzin

Università di Venezia Ca' Foscari

Latin deponent verbs are usually analyzed as idiosyncratic forms whose Middle morphology does not correspond to the subjacent syntactic/semantic structure (Embick 2000, Xu et al. 2007). This paper shows that, for the deponents produced after the first half of the II cent. BCE (ex. *ancillor* 'I serve', *dominor* 'I rule', *aquor* 'I go to get water'), the presence of the Middle morphology is syntactically justified. These deponents are denominals. Their event structure involves two events, a stative one, v-be^o, whose complement is the verbalized nominal element, and a dynamic one, v-do^o. The unique argument is both the HOLDER of the state and the DOER of the dynamic event. The Middle morphology allows for the identification between these two positions, as in a Middle reflexive derivation (Spathas et al. 2015).

Keywords: argument structure, deponents, latin syntax, middle, denominals

1. Latin deponent verbs

Latin deponents show a set of characteristic features that puzzled the ancient grammarians and still puzzle the modern linguists.¹ These verbs are always marked by the Middle morphology.² The Middle morphology is usually conceived of as

1. Similar verbs are present in Ancient Greek (Lavidas and Papangeli 2007), Modern Greek (Zombulou and Alexiadou 2014) and Albanian (Kallulli 2013).

2. The terms defining the morphologies are marked by a capital letter (e.g. Middle, Active). The terms defining the syntactic structures are not (e.g., passive, reflexive). For the Latin morphology (characterized by the morpheme, -r), I use the term Middle and not Passive. This morphology has a distribution that goes beyond the simple passive structures: it is present in passive, anticausative, reflexive and impersonal structures. Its distribution can be compared with the distribution of the SE pronoun in Italian (*se*) and German (*sich*) and with the distribution of the Middle morphology in Modern Hebrew and Modern Greek, see Schäfer (2008). Following

oppositional, a morphological element that transforms an active syntactic structure into something else, a passive, a reflexive, an anticausative, an impersonal. The example (1) shows the Active form of the verb *rumpo* ‘I break,’ while the example (2) shows the Middle one. The use of the Middle form entails, in this case, a modification of the syntactic structure from active to passive.

- (1) *Olli somnum ingens rumpit pavor.*
 him.DAT sleep.ACC great.NOM.SG break.3RDSG.PRS.ACT fear.NOM
 ‘A great fear breaks his sleep.’ (Verg. Aen. 7, 458)
- (2) *Qua ante rumpebar.* (Cic. ad Q. fr. 3, 7, 1)
 which.ABL.PL before break.3RDSG.IPFV.MID
 ‘From which things, before, I was broken.’

Deponent verbs, instead, always appear in the Middle form:

- (3) *Dominari in suos.* (Cic. Cato 38)
 rule.INF.PRS.MID in his.ACC.PL
 ‘To rule over your own men.’

The existence of such a category of verbs goes against the oppositional analysis of the Middle morphology: there is no Active form from which the example in (2) could be derived.³ Moreover, the subject in (3) (a PRO) is an Agent performing a controlled action of “ruling”: the syntactic structure is active. The target of this paper is to provide a formal analysis of the Middle morphology able to explain its presence on a specific subset of deponents, the denominals. A cross-linguistic note: in many Romance languages, like Italian, there are verbs that obligatorily present a *se* ‘self’ pronoun (e.g., It. *impadronirsi* ‘to seize’). These Romance verbs can be formally analyzed following the proposal put forward in this paper. The relevance of the different reflexivization strategies (Latin → morphology; Romance → *se* pronoun) and other peculiarities, as the fact that in Italian *se*-marked verbs obligatorily present a prefixed preposition (*in-*, *s-*, *ad-*), while in Latin it is not the case, have to be deepened.

Alexiadou and Doron (2012), I use the term Passive just for the morphologies that appear only with passive structures.

3. There is a subset of verbs that are considered deponents even if they have an Active counterpart, ex. *ludifcor/ludifico* ‘I make fun of someone’. The reason is that between the Active and the Middle verb there is no difference in meaning: The Middle form is not related to a patent passive, anticausative, reflexive or impersonal meaning.

2. Deponent verbs as idiosyncratic forms

The analysis of Latin deponents has been, recently, the topic of a few papers: Embick (2000), Xu et al. (2007), Gianollo (2010, 2014). Gianollo (2010, 2014) focuses on a relevant subset of deponents, the unaccusatives. These verbs do not present in their derivations an Agent, but an Undergoer; e.g., *morior* ‘I die’, *nascor* ‘I am born’, *orior* ‘I rise’, *aboriscor* ‘I vanish’. I will not analyze this subset in the present paper.

Xu et al. (2007) scrutinize the entire set of Latin deponents trying to find a common feature that could explain the presence of the Middle morphology. They analyze the verbal classes in which deponents fall and propose a lexical-semantic solution: Latin speakers functionalized the Middle morpheme as a marker of the lexical-semantic distinction between *canonically-active* and *non-canonically active* verbs, with the second ones marked by the Middle morpheme.

Embick (2000), instead, proposes to implement the analysis of Latin deponents within the Distributed Morphology (DM) framework (Halle and Marantz 1993). He formally explains the presence of the Middle morphology assuming the presence of an inherent [+pass] feature in the verbal root ($\sqrt{\text{V}}$).⁴ Deponent $\sqrt{\text{V}}$ s are idiosyncratically marked by an inherent [+pass] feature that is not projected during the syntactic derivation. There are, then, two ways in which the [+pass] feature could be present in a syntactic derivation. The speaker could actively merge it to derive a passive, reflexive, anticausative or impersonal derivation, or it could be inherently present in the feature matrix of the $\sqrt{\text{V}}$. In the second case the derivation can be syntactically active, but the morphological out-put would still be marked by the Passive (in his terms) morphology.

Xu et al. (2007), even if they focus on the entire set of deponents, identify as relevant a specific subset, the denominals. In this paper I focus only on this specific subset, proposing a principled and not idiosyncratic reason for the presence of the Middle morphology on these verbs.⁵ The next Section highlights the reason behind the choice of denominals.

4. He calls this feature [+pass(ive)]. See fn. 2.

5. In the spirit of Zombolou and Alexiadou (2014).

3. The productive deponents: Denominal verbs

Many deponent verbs present their first occurrence in the time span between the II-I cent. BCE and the IV-V cent. CE:⁶ these verbs are new formations. There is, then, in Latin, a productive process that creates new deponents. Examples of these verbs are: *ancillor* 'I serve', *dominor* 'I rule', *aquor* 'I go to get water', *glorior* 'I glorify myself', *famulor* 'I serve'. The productive deponents are all denominal/deadjectival⁷ formations. The table in (4) clarifies the issue⁸.

- (4) I. (Plautus – Ennius): non-denominals 56, denominals 67.
 II. (Terence – Cato): non-denominals 4, denominals 28.
 III. (Cicero – Ovid): non-denominals 1, denominals 32.
 IV. (Columella – Martial): non-denominals 0, denominals 15.
 V. (Apuleius – Tertullian): non-denominals 0, denominals 7.
 VI. (from the III cent. CE on): non-denominals 1, denominals 34.

In the first period (roughly from the III cent. BCE to the first decades of the II cent. BCE) I identified the first occurrence of 56 non denominal deponents and 67 denominals. In the subsequent period, the first occurrences of the non-denominals drop to 4, while the first occurrences of the denominals are 28. From Cicero on, finally, the new formations of non-denominals disappear.⁹ The few exceptions that I found are fully justified, e.g., *spernor* 'I despise' (VI period), from *asperno* 'I despise' (I period), has been created for metrical reasons.

The data show that denominal deponents are morphosyntactically productive. Behind synchronic productivity there is a generalizable synchronic grammatical rule. In the next sections I formalize this grammatical rule, showing that the presence of the Middle morphology on denominal deponents follows from a specific syntactic configuration and not from an inherent feature of the $\sqrt{\quad}$ or from idiosyncratic processes.

6. Dealing with a written literary corpus, I can provide only a *terminus ante quem* for the creation of these verbs, not an exact date.

7. Deadjectivals represent a small subset of the productive deponents, I will use the term denominals to refer to the entire set of productive deponents.

8. The data in (4) have been manually collected from the OLD (Glare 2012) and checked with an Etymological dictionary (De Vaan 2008). I did not take into account prefixation. There are many cases in which a deponent verb gets a new prefix in later stages of Latin. Prefixation usually adds a specific spatial/temporal semantics to the deponent verb. I have not taken into consideration such a kind of morphosyntactic formations as new productions (see Crocco-Galès and Iacobini 1992).

9. Except for the prefixation processes, that I did not take into account (see fn. 8).

4. Latin middle morphology, a formal analysis

4.1 The framework

In this analysis, I follow the Constructivist framework.¹⁰ More specifically, I follow the DM variant of the Constructivist approach.¹¹ The main point that differentiates this approach from others (e.g., the Lexicalist one, Levin and Rappaport-Hovav 1995) is that argument structure is conceived of as a syntactic fact: it is formed by the subsequent merging of a small set of syntactic heads. These syntactic heads are terminal nodes and, in a DM spirit, their morphological spell-out follows the Subset Principle (Halle and Marantz 1993, p. 122). There is an ontology of primitives, i.e. a set of functional event-denoting (eventive) heads. I adopt the labels proposed by Cuervo (2014, 2015):^{12,13}

- (5) a. v-be^o: stative event
 b. v-go^o: dynamic uncontrolled event
 c. v-do^o: dynamic controlled event

The lexical element ($\sqrt{\quad}$) is not conceived of as grammatically relevant for the building of the event structure (contra see Levin and Rappaport-Hovav 1995, 2005, Koontz-Garboden 2009). How is it merged, then? In a DM Constructivist spirit, I merge the lexical $\sqrt{\quad}$ s as adjuncts to the eventive heads.¹⁴ The lexical elements do

10. Hale and Keyser (2002), Borer (2005), Marantz (2005, 2013), Ramchand (2008), Harley (2009), Acedo Matellán and Mateu (2013), Cuervo (2014, 2015).

11. Marantz (2005, 2013), Harley (2009), Cuervo (2014, 2015).

12. See Wood (2015) for a different proposal. In Latin, there is morphological evidence for the existence of three eventive heads, stative $-\bar{e}$ (v-be^o), inchoative $-sc$ (v-go^o), agentive $-\bar{a}$ (v-do^o). Stative $-\bar{e}$ must not be confused with causative $-\bar{e}$. This insight needs further evidences in support, especially with respect to the fact that stative $-\bar{e}$ is not synchronically productive in Classical Latin.

13. In addition to this small set there is a second stative head, Appl^o (Pylkkänen 2008). This head realizes a possessive stative event: possession of a DP (Possessor) or possession of an event (Benefactive). This head is relevant in the analysis of the second and third kind of denominal deponents (see Section 5). Since I will not analyze them, I will not take it into consideration.

14. This mechanism is also called conflation, see Acedo-Matellán and Mateu (2013). See Harley (2014) for a different approach, in which $\sqrt{\quad}$ s are heads of specific $\sqrt{\quad}$ Ps. For an argument against the existence of $\sqrt{\quad}$ Ps, see Alexiadou (2014).

not modify the event (a “doing event” remains a “doing event”) but add their specific encyclopedic meaning. Let us see an example with It. $\sqrt{\text{corr}}$ - ‘run’:¹⁵

- (6) $v\text{-do}^\circ + \sqrt{\text{corr}}$ - (‘run’) \rightarrow a dynamic “doing” event of running

Each eventive head is related to the introduction of a specific argument: $v\text{-do}^\circ \rightarrow$ DOER, $v\text{-go}^\circ \rightarrow$ UNDERGOER, $v\text{-be}^\circ \rightarrow$ HOLDER (of the state). Following Wood and Marantz (2017), the actual syntactic introduction of the arguments is mediated by an argument-introducing head, $i^{*\circ}$.¹⁶ The characteristic features of this head:

- (7) a. It is endowed with a +D selectional feature: it selects for an XP of the kind DP.
 b. It has no categorial feature: it picks up the categorial feature of the XP it merges with.
 c. It closes off the extended projection of the XP it merges with. There can be only one argument for each head.

The standard $i^{*\circ}$ is:

- (8) $\{\lambda x \lambda e [\text{ROLE}(e, x)], +D\}$

This head provides a DP (x) with a specific ROLE related to a specific eventive head. The eventive head introduces the event, while the $i^{*\circ}$ introduces the syntactic and semantic space for the merging of the argument related to that event, there is no overlap between the two heads, they perform two distinct semantic and syntactic tasks.

15. See Harley (2014) for a DM formalization of the status of the \sqrt{s} . In her proposal, \sqrt{s} are marked by numbers and not by phonological or semantic features. For the sake of simplicity, I follow a standard graphic notation.

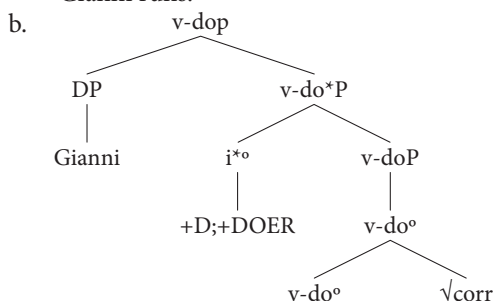
16. This head is a generalized Voice^o (Kratzer 1996). A reviewer pointed out that the system seems redundant: I project an eventive head *and* an argument-introducing head for each argument, while I could accommodate the arguments directly into the specifier position of each eventive head, eliminating the argument introducing head. Redundancy, anyway, is justified by the fact that there is an actual differentiation between the morphological output of the various eventive heads (e.g., $v\text{-go}^\circ \rightarrow \text{-sc}$, Haverling 2000) and the morphological output of the argument introducing head (the Middle morphology). If, in a DM spirit, I assume that each morphological element is related to a specific functional head, the presence of a differentiation between the eventive morphology and the Voice one leads to the conclusion that the eventive heads and the argument introducing ones are differentiated. The same reviewer notices that $i^{*\circ}$, being syncategorematic, should not present restrictions with respect to the heads it merges with. This may in fact be the case: $i^{*\circ}$ can be recategorized as a general specifier-introducing head, substituting each spec with a configuration involving a semantically relevant head and a $i^{*\circ}$ that creates the syntactic and semantic space for the merging of the spec XP. This issue remains open for further analyses.

Let us see an Italian example.

- (9) a. *Gianni corre.*

Gianni run.3RDSG.PRS

‘Gianni runs.’



The \checkmark (\checkmark_{corr} -) adjoins to *v-do°*, creating a doing event of running; *i*o* creates the syntactic and semantic space for the introduction of the DOER, the DP *Gianni*. The DP satisfies the +D requirement and gets the DOER role by functional application.

4.2 The middle morphology

This formal analysis of the Middle morphology builds on the analysis of the Voice° system proposed by Schäfer (2008, 2017) and by Spathas et al. (2015). Being *i*o* a generalized Voice° (see fn. 16), I will use the label *i*o*.

The argument-introducing head is composed by two parts, a semantic one and a syntactic one. The semantic part introduces the semantics of the argument (ROLE). This part can come in two relevant ways, standard or existential:¹⁷

- (10) a. $\{\lambda x\lambda e[\text{ROLE}(e, x)]\}$
 b. $\{\lambda e\exists x[\text{ROLE}(e, x)]\}$

In (10a) the ROLE is introduced as belonging to a specific variable that has to be saturated by functional application. In (10b), instead, the ROLE is introduced as belonging to an existentially bound element that, being existentially bound, is not a variable and does not need to be saturated by functional application. Informally, the narrow semantic component provides something like ‘there is an *x* that has the ROLE’; Since there is no open variable, there is no Full Interpretation issue. The identification of the existentially bound element with a referential element is provided by the general context. E.g., if there is a verbal derivation with a Naturally

17. See Schäfer (2017, ex. (30)). There may also be a third possibility: the semantic part could be completely expletive: no role, neither by functional application or by existential binding (Schäfer 2008). This possibility is not relevant in this context.

Reflexive $\sqrt{\quad}$ and an existentially bound DOER, the existentially bound DOER is usually contextually identified with the internal argument (reflexive interpretation); if the $\sqrt{\quad}$ is strongly agentive, instead, the DOER is usually identified with a referential element outside the derivation (passive interpretation).¹⁸

The syntactic part is the +D selectional feature. The +D selectional feature can be present and active or absent and, consequently, inactive:

- (11) a. +D
b. \emptyset

If the +D feature is active there will be a +D element (DP) merged to satisfy it. If the +D feature is inactive (\emptyset), there will be no +D element.

The definition of the Middle morphology is based on these notions.

- (12) The Middle is the morphological out-put of an argument-introducing head ($*i^{\circ}$) in which the +D feature is inactive (\emptyset).

What Latin passives, anticausatives and reflexives have in common is the presence of a syntactically inactive $i^{*\circ}$.¹⁹ The difference between these three structures relies on the type of eventive heads and on the contextual identification of the existential ROLE. Let us see, as a specimen, a reflexive derivation.²⁰

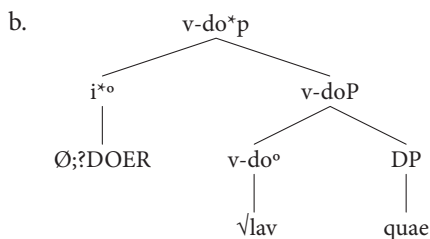
- (13) a. *Quae semper ornantur lavantur*
 who.F.PL always adorn.3RDPL.PRS.MID wash.3RDPL.PRS.MID
tergentur poliuntur. (Plaut. Poen. 217)
 scrub.3RDPL.PRS.MID smooth.3RDPL.PRS.MID
 ‘Who always adorn, wash, scrub and smooth themselves.’

18. For a complete explanation of the concept of contextual identification and of the elements that influence it, see Schäfer (2017, Section 1.4.3) and Spathas et al. (2015, Section 4.2.2).

19. The impersonal derivation behaves differently and I will not propose an analysis here. See Burzio (1986), Cinque (1988), Dobrovie-Sorin (1998), D’Alessandro (2007), Roberts (2010), Pescarini (2015).

20. For reason of space, I do not discuss anticausative examples:

- (i) *In aure quoque interdum rumpitur cartilgo* (Cels. 8, 6)
 in ear.ABL also sometimes break.3RDSG.PRS.MID cartilage.NOM
 ‘Also in the ear, sometimes, the cartilage breaks’



The DP *quae* ‘who’ is the measure of the doing event modified by $\sqrt{\text{lav}}$ ‘to wash.’²¹ The existentially bound DOER is contextually identified with *quae*. This contextual identification derives the reflexive interpretation.²² The verb is marked by the Middle morphology because the +D selectional feature is inactive (\emptyset).

5. Denominal deponents and the middle morphology

The subset of deponents under analysis, the productive one, is strongly homogeneous. First of all, as Section 3 shows, these verbs are denominal. Moreover, these verbs are all dynamic controlled change of state verbs: they all involve at least two events, a higher dynamic controlled one and a lower stative one.²³ The nominal element is always part of the stative event, in the complement position of the stative head. In the stative event the nominal element is related to a DP argument. I identified three classes of denominal deponents on the basis of the stative relation between the nominal element and the DP argument.

- (14) Identification: the argument acquires the characteristics of the nominal element, i.e. it is identified with the nominal element.

21. See Marantz (2005) for a deepening of the concept of measure DP. The measure DP is not introduced by the $v\text{-do}^\circ$. In a minimalist spirit, the $v\text{-do}^\circ$ merges with a DP, the result is a phrase labelled as the head ($v\text{-doP}$). The measure DP is not an argument of the $v\text{-do}^\circ$. A reviewer pointed out that a non-argumental element should not be able to enter in a reflexive chain. The mechanism of contextual identification, anyway, does not require argumenthood to be implemented. The only requirement is referentiality, and the measure DP is referential.

22. There are many ways in which it is possible to derive a reflexive interpretation in a natural language. In Latin, for example, there are two systems, the one sketched here, that is context based and works only with Naturally Reflexive \sqrt{s} , and the pronominal one, that uses the reflexive pronoun *se*.

23. There may be more than one stative event, as the possession+identification denominal deponents show (see below). The fact that there are no uncontrolled dynamic denominal deponents can be explained in light of the fact that there was, in Latin, a productive derivation for the uncontrolled dynamic change of state verbs: the *-sc* inchoative derivation.

Examples: *ancillor* ‘I serve’ from *ancilla* ‘servant’, *dominor* ‘I rule’ from *dominus* ‘master’, *lupor* ‘I make a prostitute of myself’ from *lupa* ‘prostitute’, *famulor* ‘I serve’ from *famulus*, ‘servant’, *circulor* ‘I make circles in order to speak’ from *circulus* ‘circle’, etc.

The actual Latin examples from the Authors can be found in the next section, in which I analyze this specific kind of deponents.

- (15) Possession: the argument ends up being in possession of the nominal element, i.e. it is the possessor of the nominal element.

Examples: *aquor* ‘I get water’ from *aqua* ‘water’, *glorior* ‘I boast’ from *gloria* ‘glory’, *annonor* ‘I get stocks’ from *annona* ‘stocked grain’, *cibor* ‘I get food’ from *cibus* ‘food’, *piscor* ‘I fish’ from *piscis* ‘fish’, *praedor* ‘I hunt’ from *praeda* ‘prey’, etc.

- (16) *Flumen erat [...] ex quo et Macedones et Romani aquabantur.* (Liv. 44, 40, 4)
 river.NOM was.3RDSG.IPFV.ACT from which.ABL and Macedonians.NOM and Romans.NOM take.water.3RDPL.IPFV.MID
 ‘There was a river [...] from which the Romans and the Macedonians used to take water.’

The meaning of this kind of verbs is clear, there is someone that acts in order to get the possession of the nominal element that forms the deponent verb.

The third kind of deponents is more complex. These denominal verbs are characterized by two obligatory arguments (the first one has Nominative case, the second one Accusative)²⁴ and a verbalized nominal element. The Accusative argument is identified with the verbalized nominal element, while the Nominative one is the beneficiary of the entire identification state.

- (17) Identification+Possession: The first argument (Accusative) acquires the characteristics (the physical location or the general characteristics) of the nominal element. The second argument (Nominative) is the beneficiary of the entire identification state.

Examples: *imaginor* ‘I picture/imagine something’ from *imago* ‘image’, *recordor* ‘I remember’ from *cor* ‘heart’, *populor* ‘I pillage’ from *populus*, ‘army’,²⁵ *testor* ‘call to testify’ from *testis* ‘witness’, etc.

24. For a proposal about the assignment of Nominative and Accusative case, see the algorithm in Schäfer (2017, Section 1.3.1).

25. Initial meaning ‘to have an army pass through something’ (De Vaan 2008).

- (18) *Vos di patrii ac penates*
 you.ACC.PL gods.VOC national.VOC.PL and penates.VOC
testor. (Cic. Sull. 31, 86)
 call.to.testify.1STSG.PRS.MID
 ‘Oh gods of my country and Penates, I call you as witnesses.’

The sentence in (18) can be paraphrased as ‘I act in order to have you (Gods and Penates) as witnesses.’ The “calling” event is not actually present; it is inferred from the general meaning: to call someone is the most direct way to have someone at your disposal.

In this paper, I will analyze only the syntax of the Identification type, reserving the other two types for further analyses.

5.1 The identification type

The semantics of these verbs could be paraphrased in an informal way as ‘x (the DP argument) acts in order to be y (the verbalized nominal element)’. Let us see some actual examples.

The identification may be with a persona or with a social role:

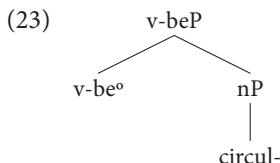
- (19) *Verum enim dotibus deleniti ultro etiam*
 indeed conversely dowties.ABL softened.PST.PTCP.NOM.PL further also
uxoribus ancillantur. (Titin. com. 70)
 wives.ABL serve.3RDPL.PRS.MID
 ‘Softened by the dowries, they (the husbands) enslave themselves for their wives even more.’
- (20) *Dominatur corpore toto.* (Lucr. 3, 276)
 rule.3RDSG.PRS.MID body.ABL entire.ABL.SG
 ‘It rules over the entire body.’
- (21) *Numquam poetor nisi si podager.* (Enn. sat. 64)
 never poetize.1STSG.PRS.MID unless if gouty.NOM.SG
 ‘I never act like a poet, unless if I’m gouty.’

The identification, anyway, is not always with an actual social role or with a human-like element, sometimes there is an identification between the argument and a geometrical form, like in *circular* ‘to speak in circles’ lit. ‘to form circles’.

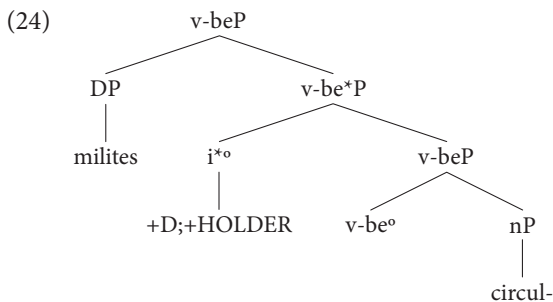
- (22) *Totis vero castris milites circulari.* (Caes. civ. 1, 64, 2)
 all.ABL.PL sed camps.ABL soldiers.NOM put.in.circle.INF.PRS.MID
 ‘In every camp, the soldiers were speaking in circles.’ (lit. ‘putting themselves in circles’)

The secondary arguments that can be present in these structures are always oblique, marked by Ablative or Dative case or by a preposition, as Ablative *corpore toto* ‘over the entire body’ in (22).

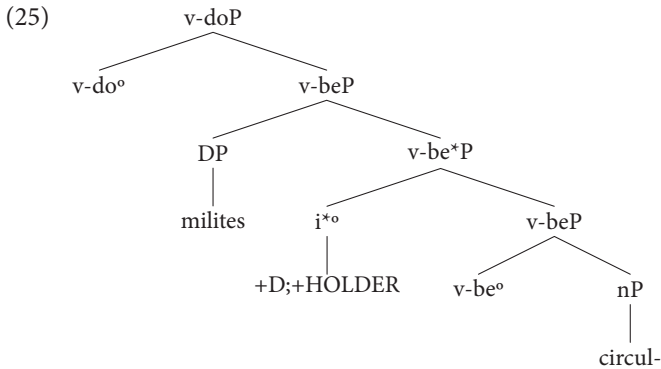
The fact that there is an identification between the argument and the verbalized nominal element directly leads to the hypothesis that in these events there is a stative part. The arguments in (19), (20), (21) and (22) are HOLDERS of specific states involving the nominal element. In (23), for example, there is a state related to the nominal element *circulus* ‘circle’:



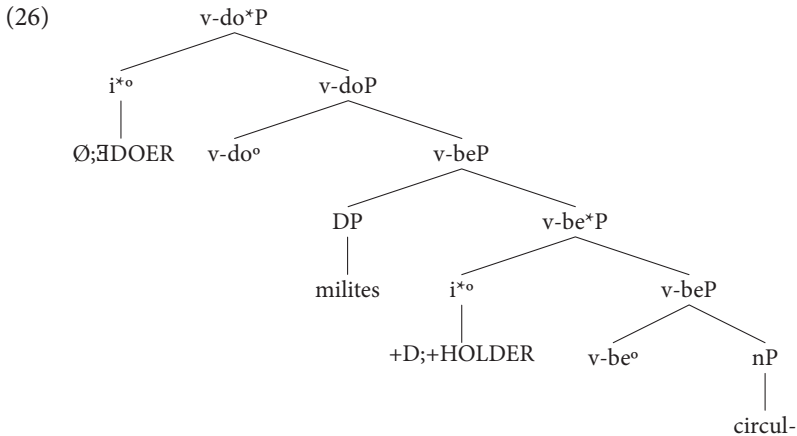
This state means ‘to be (shaped like) a circle’. *Milites* ‘soldiers’ is the HOLDER of this state.



The sentence in (24) does not only involve the described state, it also involves the dynamic event that leads to the state: the soldiers are *making* circles, they are not just “in circles”. Since this dynamic event is controlled by the argument, the event structure includes a $v\text{-do}^\circ$:



The argument related to the $v\text{-do}^\circ$, the DOER, is the HOLDER DP itself, *milites* ‘soldiers’. This DP, then, has a double set of thematic features, HOLDER and DOER. In order to reach this state of affairs, an existential and syntactically inactive $i^{*\circ}$ is merged:



The existentially bound DOER is contextually identified with the internal HOLDER and the derivation acquires its final meaning ‘the soldiers act in order to be in circles’. Following (26), the inactive $i^{*\circ}$ produces as an output the Middle morphology. The Middle, consequently, is perfectly justified and expected: there is no need for an inherent idiosyncratic feature within the feature matrix of the verbal $\sqrt{\cdot}$.²⁶

26. A reviewer pointed out that the existence of a Middle argument-introducing head is as stipulative as Embick’s (2000) inherent $\sqrt{\cdot}$ -feature. This head, anyway, provides a unitary morphological, syntactic, and semantic explanation for passives, anticausatives and reflexives in many languages (e.g., Modern Greek, Latin, Hebrew, etc.). Its stipulation covers a wide range of data, while the inherent [+pass] $\sqrt{\cdot}$ -feature explains only the deponent derivation.

6. Conclusions and open issues

Within the deponents produced after the first half of the II cent. BCE, the Middle morphology is syntactically expected and not an idiosyncratic feature of the verbal $\sqrt{\text{ }}$. It is the morphological output of an inactive existential argument introducing head that introduces into the derivation an existentially bound DOER contextually identified with the internal argument (a HOLDER in the derivations presented here). The internal argument, on the other hand, is always in a stative relation with the nominal element that gets verbalized during the derivation. Denominal deponents, by the end, are reflexives, their peculiarity, then, is not the presence of the Middle morphology, but the absence of an Active causative derivation in which the DOER is referentially distinct from the HOLDER: e.g., *ancillo* ‘I cause someone to be a maiden’. This block is entirely lexical. Being lexical, it can be easily broken. Diachronically, therefore, many denominal deponents develop an Active causative structure (e.g., causative *ancillo* in Ps. *Cypr. de singul. cleric.* 40, p. 215, 20).

The analysis of the non-denominal deponents has been avoided in this paper. Many of these verbs, e.g., *utor* ‘I fetch in my own interest’, have an internal benefactive phrase. The subject, then, is initially merged as a BENEFACTIVE and contextually identified with the \exists DOER, giving rise to the reflexive interpretation (Middle). In other cases, these verbs are former denominals. *Hortor* ‘I urge’, for example, can be analyzed as deriving from a PIE noun meaning ‘wish’ \rightarrow ‘I act so that I have someone in the wish of...’. A careful analysis of each non-denominal deponent can provide a principled explanation for the presence of the Middle morphology also for the non-denominal cases.

References

- Acedo-Matellán, Víctor, and Jaume Mateu. 2013. “Satellite Framed Latin vs. Verb Framed Romance: A syntactic approach.” *Probus* 25: 227–265.
- Alexiadou, Artemis. 2014. “Roots Don’t Take Complements.” *Theoretical Linguistics* 40 (3/4): 287–297.
- Alexiadou, Artemis, and Edit Doron. 2012. “The Syntactic Construction of Two Non-Active Voices: Passive and Middle.” *Journal of Linguistics* 48: 1–34.
<https://doi.org/10.1017/S0022226711000338>
- Borer, Hagit. 2005. *In Name Only. Structuring Sense*, vol. 1. Oxford: Oxford University Press.
<https://doi.org/10.1093/acprof:oso/9780199263905.001.0001>
- Burzio, Luigi. 1986. *Italian Syntax. A Government and Binding Approach*. Amsterdam: Springer.
- Cinque, Guglielmo. 1988. “On *si* Constructions and the Theory of Arb.” *Linguistic Inquiry* 19 (4): 521–581.

- Crocco-Galèas, Grazia, and Claudio Iacobini. 1992. "The Italian parasynthetic verbs, a particular kind of circumfix." In *Natural Morphology. Perspectives for the nineties*, ed. by Livia Tonelli, and Wolfgang U. Dressler, 127–142. Padova: Unipress.
- Cuervo, María Cristina. 2014. "Alternating Unaccusatives and the Distribution of Roots." *Lingua* 141: 48–70. <https://doi.org/10.1016/j.lingua.2013.12.001>
- Cuervo, María Cristina. 2015. "Causation Without a Cause." *Syntax* 18 (4): 388–424. <https://doi.org/10.1111/synt.12115>
- D'Alessandro, Roberta. 2007. *Impersonal si Constructions*. Berlin: Mouton de Gruyter. <https://doi.org/10.1515/9783110207514>
- De Vaan, Michiel. 2008. *Etymological dictionary of Latin and other Italic languages*. Leiden: Brill.
- Dobrovie-Sorin, Carmen. 1998. "Impersonal se Constructions in Romance and the Passivization of Unergatives." *Linguistic Inquiry* 29 (3): 399–437. <https://doi.org/10.1162/002438998553806>
- Embick, David. 2000. "Features Syntax and Categories in the Latin Perfect." *Linguistic Inquiry* 31 (2): 185–437. <https://doi.org/10.1162/002438900554343>
- Gianollo, Chiara. 2010. "I verbi deponenti latini e l'unità della flessione in -r." *Incontri triestini di filologia classica* 8: 23–49.
- Gianollo, Chiara. 2014. "Labile Verbs in Late Latin." *Linguistics* 52 (4): 945–1002. <https://doi.org/10.1515/ling-2014-0013>
- Glare, P. G. W. (ed.). 2012. *Oxford Latin Dictionary* (2nd edition). Oxford: Oxford University Press.
- Hale, Kenneth, and Samuel J. Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. Cambridge MA: MIT Press.
- Halle, Morris, and Alec Marantz. 1993. "Distributed Morphology and the Pieces of Inflection." In *The View from Building 20, Essays in Linguistics in Honour of Sylvain Bromberger*, ed. by Kenneth Hale, and Samuel Jay Keyser, 111–176. Cambridge MA: MIT Press.
- Harley, Heidi. 2009. "The Morphology of Nominalizations and the Syntax of vP." In *Quantification, Definiteness and Nominalization*, ed. by Monika Rathert, and Anastasia Giannadikou. 320–342. New York: Oxford University Press.
- Harley, Heidi. 2014. "On the Identity of Roots." *Theoretical Linguistics* 40 (3/4): 225–276.
- Haverling, Gerd. 2000. *On sco-verbs, prefixes and semantic functions. A Study in the Development of Prefixed and Unprefixed Verbs from Early to Late Latin*. Göteborg: Acta Universitatis Gothoburgensis.
- Kallulli, Dalina. 2013. "(Non-)canonical Passives and Reflexives: Deponents and their Like." In *Non-canonical Passives*, ed. by Artemis Alexiadou, and Florian Schäfer, 337–358. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.205.16kal>
- Koontz-Garboden, Andrew. 2009. "Anticausativization." *Natural Language and Linguistic Theory* 27 (1): 77–138. <https://doi.org/10.1007/s11049-008-9058-9>
- Kratzer, Angelika. 1996. "Severing the External Argument from its Verb." In *Phrase Structure and the Lexicon*, ed. by Johan Rooryck, and Laurie Zaring, 109–137. Dordrecht: Springer. https://doi.org/10.1007/978-94-015-8617-7_5
- Lavidas, Nikolaos, and Dimitra Papangeli. 2007. "Deponency in the diachrony of Greek." In *Deponency and Morphological Mismatches*, ed. by Matthew Baerman, Greville G. Corbett, Dunstan Brown, and Andrew Hippisley, 97–126. Oxford: Oxford University Press. <https://doi.org/10.5871/bacad/9780197264102.003.0005>
- Levin, Beth, and Malka Rappaport-Hovav. 1995. *Unaccusativity: At the Syntax-lexical Semantics Interface*. Cambridge, MA: MIT Press.

- Marantz, Alec. 2005. "Objects out of the Lexicon! Argument-Structure in the Syntax." Ms., MIT.
- Marantz, Alec. 2013. "Verbal Argument Structure – Events and Participants." *Lingua* 130: 152–168. <https://doi.org/10.1016/j.lingua.2012.10.012>
- Pescarini, Diego. 2015. *Le costruzioni con si. Italiano, dialetti, lingue romanze*. Roma: Carocci.
- Pylkkänen, Liina. 2008. *Introducing Arguments*. Cambridge MA: MIT Press.
<https://doi.org/10.7551/mitpress/9780262162548.001.0001>
- Ramchand, Gillian Catriona. 2008. *Verb Meaning and the Lexicon. A First-Phase Syntax*. Cambridge: Cambridge University Press.
- Roberts, Ian. 2010. *Agreement and Head Movement*. Cambridge MA: MIT Press.
<https://doi.org/10.7551/mitpress/9780262014304.001.0001>
- Schäfer, Florian. 2008. *The Syntax of (Anti)-Causatives- External Arguments in Change of state contexts*, Amsterdam: John Benjamins. <https://doi.org/10.1075/la.126>
- Schäfer, Florian. 2017. "Romance and Greek Medio-passives and the Typology of Voice." In *The Verbal Domain*, ed. by Roberta D'Alessandro, Irene Franco, and Ángel Gallego, 129–152. Oxford: Oxford University Press.
- Spathas, Giorgos, Artemis Alexiadou, and Florian Schäfer. 2015. "Middle Voice and Reflexive Interpretations: Afto-prefixation in Greek." *Natural Language and Linguistic Theory* 33: 1293–1350. <https://doi.org/10.1007/s11049-014-9279-z>
- Wood, Jim. 2015. "How Roots Do and Don't Constrain the Interpretation of Voice." Presentation at Roots IV, NYU, 29 Jun – 2 Jul 2015.
- Wood, Jim, and Alec Marantz. 2017. "The Interpretation of External Arguments" in *The Verbal Domain*, ed. by Roberta D'Alessandro, Irene Franco, and Ángel Gallego, 255–278. Oxford: Oxford University Press.
- Xu, Zheng, Mark Aronoff, and Frank Anshen. 2007. "Deponency in Latin." In *Deponency and Morphological Mismatches*, ed. by Matthew Baerman, Greville G. Corbett, Dunstan Brown, and Andrew Hippisley, 127–144. Oxford: Oxford University Press.
<https://doi.org/10.5871/bacad/9780197264102.003.0006>
- Zombolou, Katerina, and Artemis Alexiadou. 2014. "The Canonical Function of Deponent Verbs in Modern Greek." In *Morphology and Meaning. Selected papers from the 15th International Morphology Meeting*, ed. by Franz Rainer, Wolfgang U. Dressler, Francesco Gardani, and Hans Christian Luschützky, 331–344. Amsterdam: John Benjamins.

Against control by implicit passive agents

Lisa A. Reed

The Pennsylvania State University

Landau (2010, 2013) and van Urk (2011, 2013) argue that the understood agent of a passive verb is syntactically projected as a weak implicit argument. As such, it participates actively in Agree and predication, the mechanisms they assume are responsible for antecedent determination in control structures. This article examines French data involving control and passivization and proposes an alternative explanation for the facts, one that makes the diametrically opposed assumptions that the implicit agent of a passive verb is syntactically unprojected and that the reference of PRO is determined post-syntactically.

Keywords: control, implicit arguments, passive, Visser's Generalization, logophoricity

1. Introduction

Landau (2010, 2013) and van Urk (2013) examine the interaction of control and passivization in Germanic sentences parallel to the French personal and impersonal passives in (1) and (2) in order to argue for their hypotheses that (a) the understood thematic subject of a passive verb is syntactically projected as a weak implicit argument (WIA) and (b) the reference of PRO in contexts of obligatory control is primarily determined by the syntactic processes of Agree and predication.

- (1) *Le jeu a été joué [en PRO tirant le levier sur le côté de la machine].*
the game has been played while pulling the lever on the side of the machine]

machine

'The (slot) game was played by pulling the lever on the side of the machine.'

- (2) *Il a été proposé/décidé [de PRO fonder une nouvelle nation].*
it has been proposed/decided of to found a new nation

'It was proposed/decided to establish a new nation.'

The goal of this article is to show that a more in-depth examination of the facts points in the opposite direction, i.e., in favor of the more traditional views that (a) the implicit thematic subject of a passive verb is unrepresented in the syntax, as originally argued in Government-Binding terms in, e.g., Williams (1985, 1987), Rizzi (1986), and Roeper (1987) and, in Minimalist terms, in Bruening (2013) and Saab (2014) and (b) the reference of PRO is determined post-syntactically, as originally suggested in transformational terms in Chomsky and Lasnik (1977), and, in Minimalist terms, in Reed (2014: Chapter 7).

The discussion proceeds as follows. Section 2 lays out some aspects of the theories of control and passivization advocated by Landau, van Urk, and myself, the purpose being not to thoroughly compare and contrast our analyses, but to make it clear where we are in agreement and where we are not. Section 3 then compares the two approaches in terms of empirical coverage. The final section summarizes the results.

2. Landau (2000, 2010, 2013), van Urk (2011, 2013) and Reed (2014): Some points of agreement and contention

The analyses of control that Landau, van Urk, and I have developed share a number of features that make our theories much more similar to each other than they are to alternative approaches, such as those involving NP-movement, inspired by early work in O'Neil (1995) and Hornstein (1999), or implicit arguments, building off seminal ideas in Montague (1974: Chapter 8) and Brame (1978). More specifically, we agree (a) that the subject of a control clause (PRO) is licensed only by a particular feature bundle (b) that the computational system determines the respective distribution of obligatorily (OC) and non-obligatorily controlled (NOC) PRO, and (c) that PRO_{OC}'s reference is determined by an interaction of syntactic and semantic factors, whereas PRO_{NOC}'s reference is determined by pragmatic factors that include logophoricity, topicality, and genericity. In this section, these commonalities will be considered in turn, as will some important differences that emerge when the two approaches are more closely compared.

To begin, Landau, van Urk, and I follow Chomsky and Lasnik's (1995) seminal Minimalist work on control by assuming that the subject of a control clause (PRO) is licensed by a specific feature bundle. Landau (2013) and van Urk (2013) follow Landau (2000) in assuming that PRO is licensed by any T that is not associated with interpretable [+T, +Agr] features. That is, when T is associated with an interpretable semantic tense (+T) feature and is overtly inflected for person, number, and gender (interpretable +Agr) features, then it is assumed to also be associated with an uninterpretable [+R] (referential) feature that PRO cannot check

off since it lacks inherent phi-features (and so is associated with an interpretable $[-R]$ feature).

For example, none of the embedded Ts in (1) and (2) above are associated with the “fatal” $[+T(i), +Agr(i), +R(u)]$ feature bundle. They clearly are not since the embedded Ts are not overtly inflected for phi-features, which means T has a $[-R(u)]$ feature that $[-R(i)]$ PRO can check off (in more recent terms, value).

In Reed (2014: Chapter 6), I also develop a feature-based approach to the licensing of PRO. However, I follow Bowers (2002) in assuming that what licenses PRO is the absence of a functional head with unvalued phi-features that would have to undergo Agree with an inherently phi-associated nominal in order for the derivation to converge, which, I agree with Landau and van Urk, PRO is not. Thus, PRO is licensed in (1)–(2) because the embedded T is phi-incomplete, i.e. does not possess phi-features that could only be valued via Agree with an inherently phi-associated NP, such as *they*.

A second area of agreement involves the relative distribution of PRO_{OC} and PRO_{NOC} . We agree that PRO_{OC} is the “default,” found in any syntactically transparent context, whereas PRO_{NOC} is attested “Elsewhere,” i.e. in syntactically opaque contexts or in structurally ambiguous contexts in which OC results in semantic anomaly. Importantly, however, what constitutes a transparent or an opaque domain differs under the two approaches.

Landau and van Urk assume that an embedded ForceP (a.k.a. CP) does not create an opaque domain. They are firmly committed to this view because it allows them to accommodate Partial Control (PC) in examples like (3). Specifically, the head of ForceP is assumed to be unspecified for semantic number. Therefore, when it mediates the control relation, PRO is able to differ from its controller in this respect.

- (3) *Professor Smith_i T⁰ wants* $[_{ForceP/CP} Force^0 [PRO_{i+} to [(PRO) meet after class]]]$.
- $\boxed{\hspace{2cm}}$ $\boxed{\hspace{2cm}}$ $\boxed{\hspace{2cm}}$ $\boxed{\hspace{2cm}}$
 Agree Agree Agree Agree

Landau (2004: 848)

(cf. *Professor Smith met after class.)

As is clear from my BOC in (4), I adopt the opposing view that ForceP complementation correlates with NOC or, put differently, that FinP complementation results in OC.

- (4) By default, PRO must take as its antecedent a c-commanding implicit or explicit argument within the superordinate clausal domain that immediately dominates the clause in which it appears, with lexical specifications ruling out potential antecedents and ForceP constituting a phase that “closes off”

the search space. If there are no *c*-commanding potential antecedents or the search space is closed off by ForceP, PRO is assigned the index *arb* and its phi-feature specifications are logophorically determined. Reed (2014: 302)

To accommodate PC, I first associate examples like (3) with structures of the type in (5) in which the embedded FinP correlates with OC.

- (5) *Professor Smith wants* [_{FinP} [_{TP} to [_{vP} PRO *meet after class*]]].

I then follow Jackendoff and Culicover (2003: 548–551) in assuming that the semantic anomaly of control by the singular matrix subject results in conventionalized coercion of an implicit group argument that includes that agent into the argument structure of *want*. The BOC in (4) then selects that implicit argument to serve as the controller of PRO. In other words, (4) incorporates Williams' (1985: 303, 311–312) hypothesis that certain implicit arguments, although syntactically unprojected, are nonetheless accessible to the theories of binding and control. His sentence in (6) shows, e.g., that the implicit object of active voice *promise* (represented by the index on the verb) triggers a Condition C violation, it being unable to refer to the same individual as the R-expression *the doctor*.

- (6) **Mary went to the doctor's office, and she promised_k that the doctor_k would not see her again until she was really sick.*

Williams accounts for the violation by proposing that the *c*-command domain of such arguments corresponds to that of the head with which they are associated. In other words, the implicit object in (6) *c*-commands any category that the verb *promise* *c*-commands. A parallel line of reasoning accounts for the PC reading of (5): *wants*, and so also its implicit object, *c*-commands PRO, allowing the object to serve as controller.

There is a final area in which we have reached consensus. Namely, we assume, in line with work that dates back to Postal (1970), that the reference of PRO_{OC} is determined by an interaction of syntactic and semantic factors, while that of PRO_{NOC} is determined by pragmatic factors that include logophoricity, topicality, and genericity. Just logophoricity will play an important role in the discussion to follow, so only it will be considered here.

The term “logophoricity,” first coined in Hagège (1974), refers to the fact that in many languages there are overt pronouns that must refer to the individual whose perspective is being communicated whenever they are embedded under a verb of saying, knowing, thinking, perceiving, or showing emotion. So while English *she* in a sentence like *Mary said that she left* may or may not be understood to refer to Mary, its equivalent in Aghem must refer to the matrix subject, as made clear in (7). In other words, in Aghem, logophoricity is grammaticalized.

- (7) *Nnsin_y dze enyia é_{y/*x} bv# nù.*
 Nsen say that she.LOG fall.FOC
 ‘Nsen_y said that she_{y/*x} fell.’

Butler (2009: 2)

Although French and English lack a set of overt logophoric pronouns, Williams (1992) observed that PRO_{NOC} can be logophoric since it is frequently only judged grammatical if it refers to the logophoric center of the sentence – to the person whose thoughts, feelings, or speech is being reported. His data in (8a–c) make this point clearly. In (8a), Bill is obviously the logophoric center and native speakers report that PRO must be interpreted as referring to him. In (8b), the same is true of the narrator of the story. Finally, (8c) is unacceptable because it lacks a logophoric center, making it pragmatically impossible to establish PRO_{NOC}’s reference. Some additional well-known logophoric effects can be seen in (9). Such data establish that, although PRO_{NOC} does not have its reference syntactically determined, it is still often subject to obligatory pragmatic control by the logophoric center.

- (8) a. [PRO_{NOC=x} *having just arrived in town*], *the main hotel seemed to Bill_x to be the best place to stay.*
 b. [PRO_{NOC} *having just arrived in town*], *the main hotel was a vision indeed.*
 c. * [PRO_{NOC} *having just arrived in town*], *the main hotel collapsed on Bill.*
- (9) *Cindy told Tim [that [PRO_{NOC} undressing herself / himself / themselves / *ourselves in public] was a very bad idea.*

Let us conclude this section with a second area in which Landau, van Urk and I do not agree – namely, in how we view the syntactic status of the understood agent of a passive verb. While Landau and van Urk argue that such agents are syntactically projected, in the fashion indicated in bold in (10a), I assume that they are not, as indicated by the lack of IMP in (10b). I actually go even further and follow Parsons (1990), Lasersohn (1993), Bruening (2013), and Saab (2014) in assuming that the passive morphology essentially suppresses this argument from a verb’s argument structure, it only remaining interpretively available either because of a semantic rule of existential closure, as Bruening (2013: 23) and Saab (2014: 139) suggest, or because of Meaning Postulates (MPs) of the type in (11), as Lasersohn (1993: 159) has argued.

- (10) a. *Il a été **IMP** décidé [de PRO fonder une nouvelle nation].*
 it has been decided of to.find a new nation
 ‘It was decided to establish a new nation.’
 b. *Il a été décidé [de PRO fonder une nouvelle nation].*
- (11) $\forall e[\text{ATOM}(\mathbf{decide}, e) \rightarrow \exists x \text{AGENT}(x, e)]$

In short, I depart from Landau and van Urk in assuming that the understood subject of a passive is never syntactically active at LF and, therefore, cannot ever serve as a syntactically designated controller. In this respect, therefore, it differs from the type of implicit objects attested in active voice examples like (6).

Having considered where our approaches converge and diverge, let us turn now to areas in which they make opposing empirical predictions.

3. Further examining the interaction of control with passivization

3.1 Reconsidering Visser's effects

Evers (1975) was the first to observe, with respect to Dutch, that subject control verbs can undergo impersonal passivization, making them "exempt" from Visser's Generalization "Subject control verbs do not allow passivization." (12a,b) show this is equally true in French.

- (12) a. *Il a été décidé* [*de* PRO *fonder une nouvelle nation*].
 it has been decided of to. found a new nation
 'It was decided to establish a new nation.'
- b. **Paul a été menacé* [*d'*PRO *intenter un procès contre lui*].
 Paul has been threatened of to. file a lawsuit against him
 *'Paul was threatened [PRO to initiate legal proceedings against him].'

Landau and van Urk account for this fact in terms of Agree. Namely, they associate (12a,b) with the structures in (13a,b). (13a) is grammatical because the matrix T does not undergo Agree with the expletive, that NP only undergoing merge to satisfy the EPP. This means T is free to agree with the WIA and, since that argument is the semantically designated controller, the result is well-formed. In the personal passive in (13b), however, T must agree with the NP in its Spec in order for that nominal's case requirements to be met. This means that NP is the syntactically designated controller, resulting in semantic ill-formedness. In short, this account of these contrasts relies on the assumption that impersonal and personal passives both involve OC.

- (13) a. *Il a été* IMP_x *décidé* [_{ForceP} *de* PRO_x *fonder une nouvelle nation*].
 it has been decided of to. found a new nation
 'It was decided to establish a new nation.'

- b. **Paul_y, a été IMP_x menacé* [_{ForceP} *d'PRO_y intenter un procès*
 Paul has been threatened of to.file a lawsuit
contre lui].
 against him
 *'Paul_y was IMP_x threatened [_{ForceP} PRO_y to initiate legal proceedings
 against him].'

I account for these facts differently, associating them with the structures in (14):

- (14) a. *Il a été proposé/décidé* [_{ForceP} *de PRO_{NOC} fonder une nouvelle*
 it has been proposed/decided of to(found a new
nation].
 nation
 'It was proposed/decided [_{ForceP} to PRO_{NOC} establish a new nation]'
- b. **Paul_y, a été menacé* [_{FinP} *d'PRO_y intenter un procès contre lui*].
 Paul has been threatened of to.file a lawsuit against him
 *'Paul_y was threatened [_{FinP} PRO_y to initiate legal proceedings against
 him].'

(14a) is grammatical because the BOC in (4) specifies that ForceP precludes any *c*-commanding NP from serving as PRO's antecedent. NOC results and the reference of PRO is determined by logophoricity. Verbs equivalent to *propose* and *decide* designate their implicit agents as logophoric centers, so PRO is understood to refer to them.

(14b) is ungrammatical because (4) designates FinP complements as contexts of OC. Therefore, PRO must take as its antecedent the only *c*-commanding NP that is available, *Paul*, and semantic ill-formedness results. In short, I agree that personal passives involve OC, but take issue with the claim that impersonal ones do as well.

Fortunately, it is possible to determine which approach is correct. If impersonal passives involve NOC, then it should be possible to construct "atypical" pragmatic contexts in which the logophoric center cannot serve as the controller and, in these cases, one should find the three "signature" characteristics of NOC. Namely, impersonal passives should allow arbitrary and long distance control and they should license strict readings under VP-ellipsis. This is indeed the case, as established by (15)–(19). The impersonal passive in (15) has an arbitrary reading and contrasts, in this respect, with PRO_{OC} in (16).

- (15) *I contacted the selection committee about how to submit my photo. Turns out that it's preferred (by them_x) [to PRO_{arb} submit in jpeg].*
- (16) *The committee_x prefers [to PRO_{x/*arb} submit photos in jpeg].*

Example (17) shows that these constructions allow long distance control, contrasting with PRO_{OC} in (18) in this respect as well.

- (17) *Il n'a évidemment pas été décidé par les colons_x [de PRO_{NOC=y} it NEG.has obviously not been decided by the colonists of les_x taxer à ce point]. C'est bien sûr la royauté qui en a décidé them to.tax to that point it is well sure the royalty that of.it has decided ainsi.*

so

'It was obviously not decided by the colonists_x [to PRO_{NOC=y} tax them_x at such a rate]. It was the Crown_y.'

- (18) *The colonists_x obviously did not decide [to PRO_x tax *them_x / themselves_x at such a rate]. It was the Crown.*

Finally, (19) has a strict reading that the OC example in (20) lacks.

- (19) *It was proposed by Donald Trump_x [to PRO_x be the 2016 Republican candidate] and it was by the Republican National Committee too!*
(= It was proposed by the RNC that Trump be their candidate in 2016.)

- (20) **Donald Trump_x proposed [to PRO_x be the 2016 Republican candidate] and the Republican National Committee did too!*

My reviewers have pointed out an additional area that may further support my analysis. Namely, as Saab (2014: 170) observes, in Spanish *se* passive sentences like (21), the matrix T bears active voice morphology and overtly agrees in number with the embedded object. Given this, the Landau/van Urk approach would incorrectly lead us to expect such examples to be ungrammatical, due to OC by the object. However, they are fully grammatical and the implicit agent is understood to function as the controller.

- (21) *Se han decidido fundar varias compañías en el Caribe.*
CL have decided to.create several companies in the Caribbean
'They decided to establish several companies in the Caribbean.'

My approach accommodates such data under certain syntactic assumptions that may or may not prove to be correct, pending further research. Namely, if one assumes that the embedded clause takes the form of FinP and that the implicit argument of the matrix verb is accessible to the BOC in (4), then grammatical OC by that argument is expected.

In sum, Visser's related data appear to support the BOC approach. However, it does rely on arbitrary differences in c-selection. And while it would be theoretically preferable not to make use of such specifications, they appear, in this case, to

be justified, a conclusion that I, and other researchers, have reached with respect to other constructions as well. For example, to what deeper principle(s) could we attribute the fact that while causative *make* in English c-selects for ECM/small clause complements (and, as a consequence, allows preverbal embedded subjects like the one in (22)), its counterpart in modern standard French selects for “larger” complement clauses (perhaps ForceP), and therefore, disallows this word order, cf. (23a,b)?

(22) *I made [John wait].*

(23) a. *J’ai fait [patienter Jean].*

I have made to.wait Jean

b. **J’ai fait [Jean patienter].*

Such an account, it is important to note, would also have accommodate the fact that, diachronically, standard French did allow the ECM construction in (23b) and, in fact, it is still attested in certain dialects of Canadian French, cf. Reed (1992). Similarly, c-selection has been invoked to account for the syntax of modal verbs, cf. Reed (2016), the placement of Romance clitics, cf. Cardinaletti and Shlonsky (2004), and so on.

Finally, if one adopts Rizzi’s (2001) structure of the left periphery in (24), then examples like (25) indicate that impersonal passives involve an even more articulated complement than assumed thus far. That is, if Rizzi is correct in proposing that *de* ‘of’ heads FinP and that Wh-phrases move to Spec of FocP, then there is a second FocP below FinP.

(24) Force Top* Foc Top* Fin IP

(25) *C’est lors de cette étape qu’il a été décidé [ForceP [FinP de [FocP quelles*
it is during of this stage that it has been decided of which
informations retenir pour chaque lieu]]].

information to.retain for each place

‘It was during this stage that it was decided which information to retain for each place.’

3.2 A syntactic constraint on impersonal passivization

Landau and van Urk’s analysis of (13a) over-generates a large number of French and English impersonal passives:

(26) a. **Il a été refusé [de PRO discuter du problème].*

it has been refused of to.discuss of.the problem

*‘It was refused/declined to comment on the issue.’

- b. **Il a été adoré* [PRO *danser toute la nuit*].
 it has been loved to.dance all the night
 *‘It was loved to dance all night long.’
- c. **Il a été oublié* [*d’amener le vin*].
 it has been forgotten of to.bring the wine
 *‘It was forgotten to bring the wine.’

The BOC approach accounts for these facts if the matrix verbs in (26) contrast with those in (14a) in selecting for FinP. That is, (4) states that complement clauses smaller than ForceP result in OC by a c-commanding NP in the matrix. Since I assume that passive agents are entirely unrepresented at LF, they cannot ever be selected, which means that only the expletive can. However, this results in semantic ill-formedness since, as Safir (1985: 33–38) first observed in relation to examples like (27), PRO is inherently non-expletive.

- (27) a. [*Before* PRO_{NOC} *making a big decision*], *every option should be considered*.
 b. * [PRO_{NOC} *being obvious that John won’t be returning*], *we can leave*.

In short, the restricted nature of impersonal passivization in French and English further supports the view that ForceP creates an opaque domain and that the understood agent of a passive verb is not represented at LF.

3.3 On an unexpected parallel between OC PRO and overt bound pronouns

As noted above, Landau and van Urk’s account of VG makes crucial reference to the syntax of control. In (13b), for example, agreement of T with the NP in its Spec precludes T from agreeing with the implicit agent, resulting in semantic ill-formedness.

This analysis leaves unexplained the fact that pronouns like *on* ‘they’ in (28) appear to be subject to the same constraint since they too are unable to take WIAs as their antecedents.

- (28) **Paul a été IMP_x assuré* [*qu’on_x le dédommagerait*].
 Paul has been promised that one him.acc would.compensate
 *‘Paul was IMP_k promised [that they_k would compensate him].’
 (cf. Someone_k promised [that they_k would pay Paul by Friday].)

Interestingly, the implicit object of *signal* in (29a,b) is not subject to this restriction. This type of implicit argument can serve as an antecedent to an overt pronoun and PRO:

- (29) a. *Jean a fait signe* IMP_k [*qu'on_k devait partir*].
 Jean has made sign that one should to.leave
 'John signalled IMP_k [that we_k should leave].'
 b. *Jean a fait signe* IMP_k [*de PRO_k partir*].
 Jean has made sign of to.leave
 'John signalled IMP_k [PRO_k to leave].'

Finally, the implicit argument of an *impersonal* passive can serve as the antecedent to a pronoun:

- (30) *Je croyais qu'il avait été* IMP_k *décidé* [*qu'on_{k/z} se réunirait à*
 I believe that it has been decided that one ourselves to.meet in
Chicago cette semaine pour signer le contrat].
 Chicago this week for to.sign the contract
 'I believe that it was IMP_k decided/agreed/arranged [that we_{k/z} would meet
 in Chicago this week to sign this contract].'

These patterns remain mysterious under Landau and van Urk's approach. However, they are expected under a BOC analysis. First, if one assumes that the implicit agent of a passive is not represented in the syntactic structure, then one would expect it to be equally inaccessible to the theories of binding and control as both are BOCs. This, then, explains why the implicit agent cannot serve as PRO's antecedent in (14b) or as *on*'s antecedent in (28).

On the other hand, the BOC in (4) treats the implicit object associated with verbs in the active voice as being present at LF. The syntactic availability of the implicit arguments in (29a,b) accounts for the fact that they can serve as antecedents for an overt pronoun or PRO.

Finally, the grammaticality of (30) follows from the fact that the understood agent of a passive, while syntactically unavailable at LF, nonetheless remains interpretatively available via MPs or existential closure. A pronoun, therefore, can be understood to pragmatically refer to that individual, in a fashion akin to PRO in examples like (14a).

3.4 On unexpected WIA control in indirect questions

Landau (2013: 159–160) argues that indirect questions offer strong support for his approach to control and WIAs. Since the matrix subject in (31a–c) is syntactically available, it must serve as the controller of PRO, leading one to expect the Condition B violation in (31a), the impossibility of long-distance control in (31b), and the absence of a strict reading under VP-Ellipsis in (31c).

- (31) a. *John_i wondered* [_{ForceP} *who* PRO_i *to introduce himself_i/*him_i to*].
 b. **I_k thought that they_i wondered* [_{ForceP} *how* PRO_i *to feed myself_k*].
 c. *John_i remembered* [_{ForceP} *when* PRO_i *to leave*], *and Bill did too*.

(Cannot mean: Bill also remembered when John should leave.)

However, Landau (2013: 183) also makes the paradoxical observation that these constructions are problematic for his approach to VG in (13b). In (32), agreement of *was* with *Mary* should preclude *was* from agreeing with the WIA, making *Mary* the controller of PRO, but native speakers report just the opposite.

- (32) *Mary_i was* [_{VP} IMP_k *asked* [_{ForceP} *where* PRO_{*i/k} *to throw the trash*]].

The BOC approach accommodates this fuller range of data. (14b) involves FinP complementation, which the BOC identifies as a context of OC by the argument in surface subject position. However, the presence of a *wh*-element in (32) indicates that indirect questions cannot involve simple FinP complementation. It seems plausible, then, to assume ForceP complementation, which the BOC treats as a context of NOC. If so, ForceP will bar *Mary* from serving as the controller and PRO's reference will be logophorically determined, referring, as usual, to the logophoric center – the understood subject of *ask*.

As one would expect, this pragmatic control is only a tendency as it is possible to construct “atypical” contexts that preclude it, resulting in the absence of a Principle B effect in (33a), long distance control in (33b), and a strict reading under VP-Ellipsis, in (33c).

- (33) a. *Ton bébé_x ne risque pas de savoir* [*quand* PRO_{NOC=z} *le_x*
 your baby Neg is.likely not of to.know when him.acc
nourrir]. *C'est toi_z qui le sait.*
 to.feed it is you who it.acc knows
 ‘Your baby_x doesn’t know [_{ForceP} *when* to PRO_{NOC=z} *feed* him_x]. You_z do!
 b. *I_z think that my mom_x has figured out* [_{ForceP} *where* to PRO_{NOC=z+} *go on*
 our honeymoon].
 c. Speaker A: *I know I’m the only one who can do anything about this*
situation, but I just don’t know [_{ForceP} *what* to PRO *to do*].
 Speaker B: *I don’t either.*
 (Can mean I don’t know what you (= Speaker A) should do either.)

4. Against control by weak implicit passive agents

Landau and van Urk examine sentences in which passivization interacts with control to argue that the implicit agent of a passive undergoes Merge and, by virtue

of that fact, takes part in Agree and predication, which they assume determine PRO_{OC}'s reference. I have used a wider range of the same types of facts to argue in favor of the diametrically opposed views that passive agents are syntactically unrepresented and control is to be handled in terms of a BOC applying at LF.

Acknowledgements

I am indebted to Marc Authier, Dominique Sportiche, two anonymous reviewers, and the audience at LSRL 46 for helpful comments and suggestions, as well as to Francisco Ordóñez and Lori Repetti for bringing this work to print. The usual disclaimers apply.

References

- Bowers, John. 2002. "Transitivity." *Linguistic Inquiry* 33 (2): 183–224.
<https://doi.org/10.1162/002438902317406696>
- Brame, Michael. 1978. *Base generated syntax*. Seattle: Noit Amrofer.
- Bruening, Benjamin. 2013. "By phrases in passives and nominals." *Syntax* 16 (1): 1–41.
<https://doi.org/10.1111/j.1467-9612.2012.00171.x>
- Butler, Lindsay. 2009. "Explaining logophoricity, with special reference to Aghem." Ms., The University of Arizona.
- Cardinaletti, Anna, and Ur Shlonsky. 2004. "Clitic positions and restructuring in Italian." *Linguistic Inquiry* 35 (4): 519–557. <https://doi.org/10.1162/0024389042350523>
- Chomsky, Noam, and Howard Lasnik. 1977. "Filters and control." *Linguistic Inquiry* 8 (3): 425–504.
- Chomsky, Noam, and Howard Lasnik. 1995. "The theory of principles and parameters." In *The Minimalist Program*, Noam Chomsky, 13–127. Cambridge, MA: MIT Press.
- Evers, Arnold. 1975. "The transformational cycle in Dutch and German." Ph.D. dissertation, University of Utrecht.
- Hagège, Claude. 1974. Les pronoms logophoriques. *Bulletin de La Société de Linguistique de Paris* 69: 287–310.
- Hornstein, Norbert. 1999. "Movement and control." *Linguistic Inquiry* 30 (1): 69–96.
<https://doi.org/10.1162/002438999553968>
- Jackendoff, Ray, and Peter Culicover. 2003. "The semantic basis of control in English." *Language* 79 (3): 517–556. <https://doi.org/10.1353/lan.2003.0166>
- Jenkins, Lyle. 1972. "Modality in English syntax." Ph.D. dissertation, MIT.
- Landau, Idan. 2000. *Elements of Control: Structure and meaning in infinitival constructions*. Dordrecht: Kluwer Academic Publishers. <https://doi.org/10.1007/978-94-011-3943-4>
- Landau, Idan. 2004. "The scale of finiteness and the calculus of control." *Natural Language and Linguistic Theory* 22 (4): 811–877. <https://doi.org/10.1007/s11049-004-4265-5>
- Landau, Idan. 2010. "The explicit syntax of implicit arguments." *Linguistic Inquiry* 41 (3): 357–388. https://doi.org/10.1162/LING_a_00001

- Landau, Idan. 2013. *Control in generative grammar*. Cambridge: Cambridge University Press.
<https://doi.org/10.1017/CBO9781139061858>
- Lasersohn, Peter. 1993. "Lexical distributivity and implicit arguments." In *SALT III*, ed. by Utpal Lahiri and Adam Wyner, 145–161. Ithaca, NY: Cornell University.
- Montague, Richard. 1974. "The proper treatment of quantification in ordinary English." In *Formal philosophy*, ed. by Richard Thomason, 247–270. New Haven, CT: Yale University Press.
- O'Neil, John. 1995. "Out of control." In *Proceedings of the Twenty-Fifth Meeting of the North East Linguistic Society*, Volume 1 (NELS 25), ed. by Jill Beckman, 361–371. Amherst: GLSA.
- Parsons, Terence. 1990. *Events in the semantics of English*. Cambridge, MA: MIT Press.
- Postal, Paul. 1970. "On coreferential complement subject deletion." *Linguistic Inquiry* 1 (4): 439–500.
- Reed, Lisa. 1992. "Remarks on word order in causative constructions." *Linguistic Inquiry* 23 (1): 164–172.
- Reed, Lisa. 2014. *Strengthening the PRO hypothesis*. Berlin: Mouton de Gruyter.
- Reed, Lisa. 2016. "Some notes on *devoir*, *falloir*, and the theory of control." In *Romance linguistics 2013. Selected papers from the 43rd Linguistic Symposium on Romance Languages (LSRL)*, ed. by Christina Tortora, Marcel den Dikken, Ignacio L. Montoya, and Teresa O'Neill, 341–360. Amsterdam: John Benjamins.
- Rizzi, Luigi. 1986. "Null objects in Italian and the theory of *pro*." *Linguistic Inquiry* 17 (3): 501–557.
- Rizzi, Luigi. 2001. "On the position 'Int(errogative)' in the left periphery of the clause." In *Current studies in Italian syntax*, ed. by Guglielmo Cinque and Giampaolo Salvi, 287–296. Amsterdam: Elsevier.
- Roeper, Thomas. 1987. "Implicit arguments and the head-complement relation." *Linguistic Inquiry* 18 (2): 267–310.
- Saab, Andrés. 2014. "Syntax or nothing: Some theoretical and empirical remarks on implicit arguments." *Borealis: An international journal of Hispanic linguistics* 3 (2): 125–183.
<https://doi.org/10.7557/1.3.2.2952>
- Safir, Kenneth. 1985. *Syntactic chains*. Cambridge: Cambridge University Press.
- van Urk, Coppe. 2011. "Visser's Generalization: A window into the syntax of control." Ms., MIT.
- van Urk, Coppe. 2013. "Visser's Generalization: The syntax of control and passive." *Linguistic Inquiry* 44 (1): 168–178. https://doi.org/10.1162/LING_a_00124
- Visser, Fredericus Theodoricus. 1973. *A historical syntax of the English language*. Vol. III.2. Leiden: Brill.
- Williams, Edwin. 1985. "PRO and subject of NP." *Natural Language and Linguistic Theory* 3 (3): 297–315. <https://doi.org/10.1007/BF00154265>
- Williams, Edwin. 1987. "Implicit arguments, the binding theory and control." *Natural Language and Linguistic Theory* 5 (2): 151–180. <https://doi.org/10.1007/BF00166581>
- Williams, Edwin. 1992. "Adjunct control." In *Control and grammar*, ed. by Richard Larson, Sabine Iatridou, Utpal Lahiri, and James Higginbotham, 297–322. Dordrecht: Kluwer Academic Publishers. https://doi.org/10.1007/978-94-015-7959-9_9

Romance evaluative *que/che/să* sentences as inverted optatives

Cristina Sánchez López
Universidad Complutense de Madrid

Romance exclamative sentences introduced by the conjunctions *que* (Spanish, French, Portuguese) and *che* (Italian) ‘that’ and the Romanian particle *să* plus a subjunctive verb can receive an evaluative meaning and express the speaker’s displeasure or discontent about the propositional content. This paper describes and analyzes their properties. I propose that Romance *que/che/să* evaluative sentences are expressive sentences that express the speaker’s emotion about a proposition. The proposition is compared with other salient alternative propositions according to an inverted bouletic scale, so that the proposition is the less desirable for the speaker. The main formal properties (mood and tense restrictions) and semantic properties (presupposition of factivity and negative evaluation of the proposition) of the sentences naturally follow from the analysis.

Keywords: evaluative, expressive, optatives, subjunctive, force, type-clause

1. Introduction

Romance exclamative sentences introduced by the conjunction *que/che/să* ‘that’, plus a subjunctive verbal form can receive an evaluative meaning, and express the speaker’s displeasure, unease or discontent about the propositional content. The evaluative meaning can be reconstructed by predicates such as ‘I am amazed’, ‘I am shocked’, or ‘I would not have expected’:

- (1) a. It. *Che se ne sia andato da solo!* (*dopo aver fatto tutte quelle*
that REFL PART is_{subj} left of alone after have_{inf} done all those
storie)
stories
‘It is awful that he left alone, after having done all that’

(Benincà 1995: 133)

- b. P. *Que ele passe a vida a zangar-se conosco (é o cúmulo)!*
 that the spends the life to miff with us is too much
 'It is awful that he spends his life getting hungry with us! That's too
 much.' (Duarte 2003)
- c. Sp. *¡Que tenga yo un sobrino tan salvaje! clamó don Opas, dando un*
 that have_{subj} I a nephew so savage claimed Mr. O. giving a
golpe recio
 hard blow
 'It is annoying that I have such a savage nephew, Mr. O. exclaimed
 forcefully' (CORDE)
- d. Fr. *Que j'aie un neveu aussi sauvage! Incroyable!*
 that I have_{subj} a nephew so savage incredible
 'It is annoying that I have such a savage nephew!'
- e. R. *Să am eu un nepot atât de sălbatic! Ce păcat!*
 PART_{subj} have_{1SG} I a nephew so of savage be_{3SG} incredible
 'It is annoying that I have such a savage nephew!'

The existence of these kinds of statements (*que/che/să-Evaluatives* henceforth) has been noted in the literature (see Benincà 1995, Duarte 2003 and Evans 2007); however, to the best of my knowledge, a formal analysis of them has not been accomplished until now. The aim of this paper is to provide a semantic and syntactic analysis of *che/que/să-Evaluative* sentences. Concretely, I will propose that they are inverted optative sentences, that is, optative sentences that are interpreted with respect to an inverted bouletic scale. The main formal properties (related to mood and tense restrictions) and semantic properties (related to the presupposition of factivity and the negative evaluation of the proposition) of these sentences can be naturally derived from the analysis.

The structure of the paper is as follows. In Section 2, I describe the main syntactic, semantic and prosodic properties of evaluative sentences. In Section 3, I propose an analysis for *che/que/să-Evaluatives* that focuses on the expressive nature (Section 3.1), the syntactic structure (Section 3.2), the tense restrictions and the contribution of subjunctive mood (Section 3.3), and the relationship between the presupposition of factivity and the subjunctive mood (Section 3.4). Finally, conclusions are presented in Section 4.

2. Basic description of evaluative *che/que/să* romance sentences

2.1 The form of *che/que/să*-Evaluatives

Che/que/să-Evaluatives are introduced by the conjunction *que* ‘that’ in Spanish, French and Portuguese, the conjunction *che* ‘that’ in Italian, and the modal particle *să* in Romanian. The verb displays subjunctive morphology. The word order is similar to the word order in declarative sentences: the subject must precede the verb in French; in the other languages, the subject can freely precede or follow the verb.

These formal properties are shared by *che/que/să*-Evaluatives and Romance optatives headed by *che/que/să*. In all the languages considered here, the main sentences with the structure *che/que/să*+V_{subj} can receive an optative reading and express the vivid wish, hope or desire of the speaker about the propositional content. Actually, Romance main sentences with the form *che/que/să*+V_{subj} are ambiguous between the evaluative and the optative reading, as the Spanish example in (2) shows.

- (2) Sp. *¡Que trabaje él mañana!*
 that works_{subj} he tomorrow
 a. Optative reading: ‘I hope he works tomorrow!’
 b. Evaluative reading: ‘It is awful that he works tomorrow!’

Both *que/che/să*-Evaluatives and *que/che/să*-Optatives are expressive sentences that express the speaker’s emotion about a proposition. They reveal emotional or affective reactions that constitute over manifestations of emotional or affective behavior. However, they differ in the content of the emotion: optatives express the speaker’s wishes whereas evaluatives express the speaker’s negative attitude or evaluation.

In addition to their form and their expressive nature, *que/che/să*-Evaluatives and *que/che/să*-Optatives share the same temporal restrictions. They are only possible with verbs in present and present perfect tense, and reject verbs in past and past perfect tenses.

Both Spanish and Italian have four subjunctive tenses: present, present perfect, past and pluperfect; the examples in (3a) and (3b) shows that the *que/che*-Evaluative sentences are only possible with present and present perfect, but not with imperfect and pluperfect. Portuguese has three subjunctive tenses: present, past and future, the example in (3c) shows that only the present is possible. Romanian only has two subjunctive forms: present is possible, past is not, as shown in (3d). Finally, French admits both present and present perfect; imperfect and pluperfect are archaic forms, indeed excluded from this context (3e):

- (3) a. It. *Che io {abbia / abbia avuto / *avessi / *avessi avuto} un nipote*
 that I have_{subj} / have_{subj} had / had_{subj} / had_{subj} had a nephew
così selvaggio!
 so savage
- b. Sp. *¡Que yo {tenga / haya tenido / *tuviera / *hubiera tenido} un*
 that I have_{subj} / have_{subj} had / had_{subj} / had_{subj} had a
sobrino tan salvaje!
 nephew so savage
- c. P. *Que eu {tenha / *tivesse / *tiver} um sobrinho tão selvagem!*
 that I have_{subj} / had_{subj} / had_{subj.fut} a nephew so savage
- d. R. *Să {am / *fi avut} eu un nepot atât de sălbatic! Ce*
 PART have_{subj} / PART_{PAST.SUBJ} had I a nephew so of savage It's
păcat!
 awful
- e. Fr. *Que j' {aie / aie eu / *eusse / *eusse eu} un neveu aussi*
 that I have_{subj} have_{subj} had / had_{subj} / had_{subj} had a nephew so
sauvage! Incroyable!
 savage incredible

The examples in (4) show that *que/che/să*-Optatives display the same temporal restrictions, since they are not grammatical with a verb in a past tense:

- (4) a. Sp. *¡Que {tenga / *tuviera} la suerte de encontrar algo decente!*
 b. It. *Che {abbia / *avesse} la fortuna di trovare qualcosa di buono!*
 c. P. *Que {tenha / *tivesse} sorte de achar algo decente!*
 d. Fr. *Qu' il {aie / *eusse} la chance de trouver quelque chose de décent!*
 that he has_{subj} had_{subj} the luck of find_{inf} something of decent
- e. R. *Să {ai / *fi avut} noroc să găsim ceva*
 PART_{SUBJ} have_{2SG} PART_{PAST.SUBJ} had luck PART_{SUBJ} find_{2SG} something
decent!
 decent
 'I hope that you are lucky and find something good!'

Tense restrictions exclude a possible analysis of *que/che/să*-Evaluatives like truncated subordinate clauses, that is, like subordinate clauses that are a complement of a silent or elided main evaluative predicate. This kind of analysis is proposed by Benincà (1995: 137), who affirms that the kind of exclamative exemplified by (1a) is *dipendente ellittica della principale*, or, 'dependent on the elided main clause' (see also Rivero 1977 and Spaulding 1934). Subordinate sentences depending on an evaluative attitude verb admit past tenses, whereas *che/que/să*-Evaluatives do not:

- (5) a. Sp. *Era molesto que yo {tuviera / hubiera tenido} un sobrino tan*
 was_{3SG} irritating that I had_{subj} / had_{subj} had a nephew so
salvaje
 savage
 ‘It was irritating that I had such a savage nephew.’
- b. P. *Era terrível que ele tivesse um sobrinho tão selvagem!*
 was_{3SG} terrible that he had_{subj} a nephew so savage
 ‘It was terrible that I had such a savage nephew.’

The main predicate deletion hypothesis wrongly predicts that the examples in (3) and (4) ought to be grammatical with a past tense, unless a specific condition on tenses restricts the ellipsis of the matrix predicate when the subordinate verb displays past tense. As is obvious, such a condition would not be far from being an *ad hoc* stipulation.

2.2 Semantics and intonation of *que/che/sã*-Evaluatives

Que/che/sã-Evaluatives express the displeasure, unease or discontent of the speaker about the propositional content; in other words, the propositional content denoted by the utterance is evaluated negatively on the part of the speaker.¹ *Que/che/sã*-Evaluatives can be followed by statements that make explicit the negative evaluation, and reject those that express a positive evaluation, as shown in (6):

- (6) Sp. *¡Que tenga yo un sobrino tan salvaje! ¡{Terrible/#excelente}!*
 that have_{subj} I a nephew so savage terrible / excellent
 ‘That I have such a savage nephew! That’s {terrible/excellent}!’

The negative evaluation, which is only one of the wide range of evaluative attitudes that can be expressed by the speaker, crucially characterizes these types of statements. This property provides an extra argument for not considering *que/che/sã*-Evaluatives as subordinated sentences depending on a silent or elided evaluative verb. A subordinate clause introduced by *que/che/sã* with a subjunctive verb can be the argument of both predicates expressing a positive evaluation, as in (7a), and predicates expressing a negative evaluation, as in (7b). *Que/che/sã*-Evaluatives

1. Benincà (1995: 133) describes the meaning of (1a) above in the following way: “the exclamative utterance introduced by *che* ‘that’ can be interpreted like the argument of a predicate such as *è proprio inaudito* ‘it is incredible’, *è proprio il colmo* ‘it is an annoyance’”, (“*La frase [esclamativa introdotta da che] può essere interpretata come argomento di un predicato del tipo è proprio inaudito, è proprio il colmo*” – my translation). A similar description is provided by Duarte (2003) regarding (1b) above. The TLFi (*Trésor de la Langue Française, s.v. que*) states that the conjunction *que* ‘that’ followed by a subjunctive verb expresses wish, anger, surprise (“*suivi d’un verbe au subj., exprimant le souhait, l’indignation, la surprise*”, underlined is my translation).

only receive the second interpretation, which must not be the case if they were the result of the elision of the main predicate, since nothing prevents positive evaluation predicates from being elided under the same conditions that negative evaluation predicates are:

- (7) Sp. *¡Es {maravilloso / bueno} que tenga yo este sobrino!*
 is wonderful good that had_{subj} I that nephew
 ‘It is wonderful / good having such a nephew!’
 Sp. *¡Es {terrible / malo} que tenga yo este sobrino!*
 is terrible bad that have_{subj} I that nephew
 ‘It is terrible / bad having such a nephew’

The contrast between (6) and (7) allows us to reject a potential main predicate elision analysis for *que/che/sã*-Evaluatives and to conclude that they are not truncated subordinated sentences but main sentences with a complex left periphery by which the speaker expresses his/her discontent or unease.

Que/che/sã-Evaluatives are compatible with the presupposition that the proposition is a real fact. That’s why they are incompatible with a context that denies the factual presupposition, as in (8).

- (8) Sp. *¡Que tenga yo un sobrino tan salvaje! #Y de hecho no lo tengo*
 that had_{subj} I a nephew so savage and actually not him_{ac} have
 ‘It is annoying that I have such a savage nephew, and actually I don’t have one.’

Que/che/sã-evaluative sentences express the actual situation that is contrary to the expectations of the speaker. Since the proposition can denote a fact that contradicts the expectations of the speaker, the propositional content is evaluated negatively and the utterance expresses the speaker’s displeasure, unease or discontent.

This property creates a main difference between *que/che/sã*-Evaluatives and *que/che/sã*-Optatives, since all optative sentences convey the presupposition that the desired situation is not a fact. Optatives express the speaker’s vivid wish, hope or desire about a situation that is presupposed not to be factual; therefore, they are not compatible with a context that denies the non-factual presupposition, as in (9):


- (9) Sp. *¡Que tengas suerte! #...y de hecho la tienes*
 that have_{subj.2sg} luck and actually it_{ac} have_{ind.2sg}
 ‘I hope you have good luck!... and actually you do.’

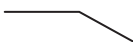
As noted by Benincà (1995), the presence of some modal elements, like deontic auxiliaries, favors the evaluative reading. The sentences in (10) are unambiguous between the evaluative and the optative reading, and only the evaluative reading is available:

- (10) a. It. *Che Mario debba comportarsi così!*
 that Mario must_{subj} behave_{inf} this way
 ‘I am amazed and shocked that Mario behaves this way.’
 (Benincà 1995: 133)
- b. Sp. *¡Que haya yo de soportar a ese idiota!*
 that have_{subj} I of stand_{inf} to that idiot
 ‘I am amazed that I have to stand that idiot.’

The presence of the modal auxiliary is a sufficient, but not a necessary, condition to interpret the utterance as evaluative. The modal auxiliary suspends the reality of the situation by putting it under the effect of a deontic operator. In example (10b), the speaker expresses his/her discontent or displeasure. However, it is important to recall that the speaker is not displeased because (s)he actually tolerates a stupid person; (s)he regrets that (s)he must do it. The interaction between desirability and obligation weakens the emotion; in other words, the evaluation of an undesirable situation makes sense when the speaker cannot avoid such a situation because of an obligation or an unavoidable necessity.

Finally, prosody marks one more difference between *que/che/sã*-Evaluatives and *que/che/sã*-Optatives. Like other expressive sentences, both *que/che/sã*-Evaluatives and *que/che/sã*-Optatives are prosodically marked by exaggerated acoustic intensity (wide ranging peaks and troughs) and orthographically indicated by exclamation marks « (¡)! », *but* their intonation can partially differ from each other. To illustrate this point, I will focus on Spanish, although similar considerations could be added for the other Romance languages. In Spanish, the difference between the two readings is marked at the end of the sentence: there is a rising intonation movement at the end of the sentence in *que*-Evaluative sentences and this pattern coincides with the lengthening of the final syllable of the sentence (11); this intonation contrasts with the falling final intonation of *que*-Optatives (12):²

- (11) 
 Sp. ¡Que trabaje él mañana!
 that works_{subj} he tomorrow
 ‘It is awful that he works tomorrow!’

- (12) 
 Sp. ¡Que trabaje él mañana!
 that works_{subj} he tomorrow
 ‘I hope he works tomorrow!’

2. The intonation patterns have been analyzed with PRAAT (Boersma and Weenink 2015). Technical data cannot be presented here for space reasons. See Sánchez López (in press) for details.

The relationship between intonation and the expression of the speaker's negative attitude in Spanish was observed and cleverly described by Navarro Tomás (1974), who wrote:

The rising inflection occurs in exclamations expressing amazement and surprise which involve, at the same time, the intention to make a reply, correction or protest. The amazement and surprise is often referred to facts, purposes or attitudes whose allocation is considered inappropriate (...). Exclamative utterances with a sense of disconformity, censure or disapproval exhibit rising intonation as well.
(Navarro Tomás 1974: 175ss, my translation)

A deeper research would be necessary in order to clarify how the link between intonation and semantics is established. I must leave this issue open here.

2.3 Summary

Que/che/să-Evaluatives are expressive sentences with a characteristic final rising intonation; they express the discontent, unease or disconformity by the part of the speaker about a situation that is presupposed to be factual. They also are introduced by a conjunction equivalent to *that* (Sp., Fr., and P. *que*; It. *che*) or a modal particle (R. *să*); the verb displays subjunctive morphology and is temporally constrained to present and present perfect (past and pluperfect are excluded). They can contain a deontic modal verb, which favors the evaluative reading of the sentence.

Que/che/să-Evaluatives are formally similar to *que/che/să*-Optatives, since both types of sentences are introduced by the same conjunctions and have a subjunctive verb with the same tense constraints. They both are expressive sentences, but Evaluatives express the negative evaluation by the speaker, whereas Optatives express the speaker's vivid hope. They differ also in both the intonation and the factivity presupposition.

3. The analysis of *que/che/să*-Evaluatives as inverted optatives

In this section, I will propose that *que/che/să*-Evaluatives are inverted optatives, that is, optative sentences interpreted according to an inverted bouletic scale. According to this analysis, the formal similarities between *que/che/să*-Evaluatives and *que/che/să*-Optatives are due to the fact that they are both basically optative statements, constructed over the same structural and semantic architecture. The semantic differences between them come from the fact that they are interpreted according to different bouletic scales: optatives are interpreted with respect to a bouletic scale and express what the speaker desires, whereas evaluatives are

interpreted with respect to an inverted bouletic scale and express what the speaker does not desire.

3.1 The expressive nature of *que/che/sã*-Evaluatives

My analysis of *que/che/sã*-Evaluatives is based on the hypothesis that expressive sentences, that is, sentences that convey the speaker's emotion about a proposition, contain a generalized exclamation operator *EX*, as proposed for exclamatives by Gutiérrez Rexach (1996, 2001), Castroviejo (2006), and for optatives by Grosz (2011) and Sánchez López (2016, 2017). *EX* selects a truth conditional statement (a proposition) p and a scale S and quantifies over scalar alternatives to p . In Evaluatives, S is ordered according to the speaker's preferences, such that Evaluatives are modalized propositions anchored to the world of the speaker's desires or preferences. The presence of *EX* explains the expressive nature of Evaluatives, since the operator combines with a proposition, and turns it into a felicity-conditional expression of an emotion. I take the formulations below from Grosz's (2011) hypothesis about optatives:

- (13) i. An utterance of *EX* (φ) conveys that the speaker at the point of utterance has an emotion ε (or at least an evaluative attitude ε) towards φ . By uttering an utterance of *EX*(φ), the speaker intends to express an emotion ε , rather than describe ε .
- ii. *EX*: For any scale S and proposition p , interpreted in relation to a context c and assignment function g , an utterance *EX*(S)(p) is felicitous if:
- $$\forall q [\text{THRESHOLD}(c) >_S q \rightarrow p >_S q]$$

That is to say: *EX* expresses an emotion that captures the fact that p is higher on a (speaker-related) scale S than all contextually relevant alternatives q below a contextual threshold, where *THRESHOLD*(c) is a function from a context into a set of worlds/a proposition that counts as high with respect to a relevant scale S .

I propose that the analysis in (13) is valid for both Evaluatives and Optatives since in both cases, S refers to a scale that models the speaker's preferences (i.e. a bouletic scale). Evaluatives and Optatives differ in the orientation of the scale: in the unmarked situation, the top of the bouletic scale is occupied by the most desirable proposition and the sentence has an optative reading; if S is inversely oriented, the top is occupied by the less desirable proposition and the sentence has an evaluative reading. Intonation marks an inverted orientation of the scale. I propose to add the following conditions to (13):

- (14) An utterance of type $EX(S)(p)$, where S is a bouletic scale that orders p and the relevant alternatives to p according to the speaker's preferences,
- i. $EX(S)(p)$ will have an optative meaning if $\forall q$
[THRESHOLD(c) $>_s q \rightarrow p >_s q$]
 - ii. $EX(S)(p)$ will have an evaluative meaning if $\forall q$
[THRESHOLD(c) $>_s q \rightarrow p <_s q$]

This account is consistent with the idea that desirability is the result of the comparison of a proposition and salient alternatives in the line of Villalta's (2007, 2008) analysis. I propose that the negative evaluation of the proposition is the result of desirability, and therefore, of the comparison of the proposition with its contextually relevant alternatives. The speaker evaluates the proposition negatively because the proposition is less desirable for him/her.

This analysis explains that Evaluatives only express a negative attitude by the part of the speaker, that is, that Evaluatives only express a part of the wide range of evaluative attitudes that would be possible for subordinated sentences depending on a main verbal predicate (recall the contrasts in (6) and (7) above). The positive evaluative attitude, indeed, coincides with the optative reading.

A piece of evidence for the proposal that *que/che/să*-Evaluatives involve an inverse bouletic scale is that the negative form of an evaluative is equivalent to the corresponding affirmative optative. As noted by Salvá (1830: 221), Spanish negative evaluatives with *que* can express desire and be equivalent to affirmative optatives introduced by the particle *ojalá* 'I wish'; examples in (15) are his:³

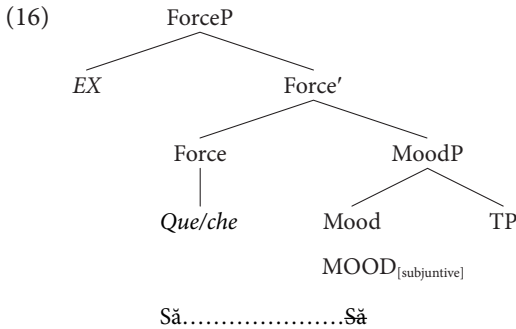
- (15) a. Sp. *¡Que no pueda yo explicar todo lo que siento!*
that not can_{subj} I explain all that feel
'It is annoying that I cannot explain what I feel!'
- b. Sp. *¡Ojalá pudiera yo explicar todo lo que siento!*
PART could_{subj} I explain all that feel
'If only I could explain what I feel!'

3.2 The syntactic structure

Associating the expressive interpretation of a sentence with an exclamatory operator entails that there has to be a syntactic projection hosting this constituent expressing force. In line with the cartographic approach, which takes CPs as

3. Although they involve an inverted scale, *que/che/să*-Evaluatives are not downward entailment contexts and do not license negative polarity items. This is consistent with the fact that factive emotive predicates that express a negative evaluation are not negative triggers and do not differ from those expressing a positive evaluation at this respect. I thank D. Sportiche for making me aware of this issue.

involving a fine-grained structure encoding topic, focus, and force constituents, I will assume that *EX* is syntactically merged in the specifier of ForceP (Rizzi 1997, Gutiérrez-Rexach 2001, Sánchez López 2016, 2017). I will propose that the conjunctions *que/che* ‘that’ merge in Force° to ensure the right clause type marking. In Romanian, the particle *să* has a double function as a mood and subordinator. Following Roussou (2010) among others, I assume *să* to be a Mood head that moves to Force.⁴ The syntactic structure, then, as shown in (16), is also valid for *que/che/să*-Optatives:



3.3 Tense restrictions and the contribution of Mood

It is not surprising that *Que/che/să*-Evaluatives, being expressive sentences interpreted according to an inverted bouletic scale, and Evaluatives and Optatives, having the same syntactic structure, exhibit the same tense restrictions. In this section, I will propose that these tense restrictions are related to the modal base the sentences are anchored to. Concretely, I will propose that tense features under subjunctive mood are linked to the modal base the sentence must be interpreted within. At the same time, the modal base is related to the different ways the ForceP layer lexicalizes.

Following Laca (2010: 198), I assume that present and present perfect subjunctive are deictic tenses, always anchored with regard to the utterance time, which provide a not completely realistic modal base (Iatridou 2000), that is, a domain that contains a word of evaluation compatible with the real world $-w_0-$ in addition to other possible words.⁵ Both *que/che/să*-Evaluatives and *que/che/să*-Optatives

4. The link between Mood head and C can be traced back to Kempchinsky's (1987) analysis of the Romance subjunctive (see also Rivero and Terzi 1995).

5. According to Iatridou (2000), a non-completely realistic modal base differs from a non-realistic modal base, which is a non-unitary domain of words of evaluating that excludes the real

are associated to a subjunctive feature in Mood and present features in Tense; they are interpreted according to a not completely realistic modal base.

Two important consequences follow from this explanation: firstly, *que/che/sã*-Optatives express wishes that are interpreted as feasible desires, which are compatible with the actual state of affairs. That explains that *que/che/sã*-Optatives are anomalous if the context determines that the desired situation is impossible, as in (17):

- (17) Sp. #*Es imposible que ella venga hoy, pero ¡que venga solo un ratito!*
 is impossible that she comes_{subj} today but that comes_{subj} only a while
 ‘It is not possible that she comes today, but I hope that she comes only for a little while.’

Secondly, *que/che/sã*-Evaluatives are compatible with the presupposition that the proposition in which the speaker expresses an evaluation denotes a fact. Since they are interpreted with respect to a not completely realistic modal base, that is a modal base containing some worlds of evaluation in addition to the actual world, the factivity presupposition is not ruled out.

3.4 The factive presupposition and the subjunctive mood

Que/che/sã-Evaluatives are similar to factive emotive predicates (Kiparsky and Kiparsky 1970) in that they presuppose the truth of the proposition (factivity) and, nevertheless, require subjunctive Mood. As is known, subjunctive under factive emotive predicates looks like *realis* subjunctive cases, and clearly do not fit the traditional understanding of subjunctive as occurring in *irrealis* contexts (Givón 1994, Mithun 1995). They are neither clearly intentional (Farkas 1985, 1992) nor non-veridical (Giannakidou 1997), nor can they be considered old information in all cases (Siegel 2009) and they are not negation/polarity-licensed (Guitart 1991).

Under the analysis proposed here, the negative evaluation of the proposition is the result of comparing the proposition with other salient alternatives according to a bouletic scale: the speaker expresses an emotion about his desires by saying that what happens is the less desirable situation in the scale of his preferences. According to this analysis, no incompatibility between the factive nature of evaluatives and the subjunctive mood is predicted. This analysis is consistent with Villalta’s (2007, 2008) hypothesis about mood selection. Villalta proposes that predicates that require the subjunctive mood introduce an ordering relation between propositions by comparing the proposition to its contextually available

word. See Sánchez López (2016, 2017) about optative sentences interpreted according to a non-realistic modal base.

alternatives; the realization of subjunctive features in Mood ensures that the evaluation of alternatives happens at the correct place in the tree.

My proposal is that Villalta's analysis can be extended to the selection of mood in main clauses: *EX* ensures an ordering relation between *p* and the salient alternatives according to a scale anchored to the speaker. The evaluative reading is obtained when the proposition is evaluated with respect to an inverted bouletic scale. Subjunctive features in Mood ensure that the evaluation of alternatives is at the right place.

4. Conclusions

I have proposed that Romance *que/che/să* evaluative sentences are expressive sentences that express the speaker's negative emotion about a proposition. Their main formal properties (mood and tense restrictions) and semantic properties (presupposition of factivity and negative evaluation of the proposition) follow from an analysis that considers them like inverted optative statement.

The description of these kinds of statements expands the inventory of Romance main expressive sentences and provides a new empirical base for the knowledge of the relationship between the left periphery of the sentences, their illocutive force, and the properties of Mood. The analysis proposed here supports the idea that the illocutive force of statements has a syntactic correlate in the merge of the Force and Mood heads.

Acknowledgements

I would like to thank my French, Italian, Portuguese and Romanian informants Laurence Rouanne, Sonia Jordana, Anna Maria Cracchione, Federico Silvagni, Emmanuela Raugel and Virginia Orpisa. I am also very thankful to an anonymous reviewer for his/her comments. All remaining errors are mine. The research underlying this work has been partially supported by a grant to the project "La exclamación y sus relaciones con otras modalidades oracionales" (FFI2012-34974).

References

- Benincà, Paola. 1995. "Il tipo esclamativo." In *Grande grammatica italiana di consultazione*, ed. by Lorenzo Renzi, Giampaolo Salvi, and Anna Cardinaletti, 127–152. Bologna: Il Mulino.
- Boersma, Paul, and David Weenink. 2015. *Praat: doing phonetics by computer [Computer program]*. Version 5.4.08. Retrieved March 24, 2015 from <http://www.praat.org/>.

- Castroviejo, Elena. 2006. "Wh-Exclamatives in Catalan." Ph.D. dissertation, Universitat de Barcelona.
- CORDE *Corpus Diacrónico del Español* [on line]. Retrieved October 1-December 1, 2015 from <http://www.rae.es>.
- Duarte, Inês. 2003. "Estrutura da frase simples e tipos de frase." In *Gramática da língua portuguesa*, ed. by Maria Helena Mateus, Ana Maria Brito, Inês Duarte, and Isabel Hub Faria, 433–506. Lisboa: Caminho.
- Evans, Nicholas. 2007. "Insubordination and its Uses." In *Finiteness. Theoretical and Empirical Foundations*, ed. by Irina Nikolaeva, 366–431. Oxford: Oxford University Press.
- Farkas, Donka F. 1985. *Intensional Descriptions and the Romance Subjunctive Mood*. New York: Garland.
- Farkas, Donka F. 1992. On the semantics of subjunctive complements. In *Romance Languages and Modern Linguistic Theory*, ed. by Paul Hirschbühler, and E. F. K. Koerner, 69–104. Amsterdam: John Benjamins. <https://doi.org/10.1075/cilt.91.07far>
- Giannakidou, Anastasia. 1997. "The Landscape of Polarity Items." Ph.D. dissertation, Rijksuniversiteit Groningen
- Givón, Talmy. 1994. "Irrealis and subjunctive." *Studies in Language* 18 (2): 265–337. <https://doi.org/10.1075/sl.18.2.02giv>
- Guitart, Jorge M. 1991. The pragmatics of Spanish mood in complements of knowledge and acquisition-of-knowledge predicates. In *Discourse-Pragmatics and the Verb*, ed. by Suzanne Fleischman, and Linda R. Waugh, 179–193. London: Routledge.
- Grosz, Patrick Georg. 2011. "On the Grammar of Optative Constructions." Ph.D. dissertation, MIT.
- Gutiérrez-Rexach, Javier. 1996. "The Semantics of Exclamatives." In *Syntax at Sunset. UCLA Working Papers in Linguistics*, ed. by Edward Garret, and Felicia Lee, 146–162. Los Angeles: UCLA.
- Gutiérrez-Rexach, Javier. 2001. "Spanish Exclamatives and the Semantics of the Left Periphery." In *Selected papers from Going Romance 99*, ed. by Johan Roorych, Yves de Hults, and Jan Schroten, 167–194. Amsterdam: John Benjamins.
- Iatridou, Sabina. 2000. "The Grammatical Ingredients of Counterfactuality." *Linguistic Inquiry* 31 (2): 231–270. <https://doi.org/10.1162/002438900554352>
- Kempchinsky, Paula. 1987. "The subjunctive disjoint reference effect." In *Studies in Romance Languages*, ed. by Carol Neidle, and Rafael A. Nuñez Cedeño, 123–140. Dordrecht: Foris. <https://doi.org/10.1515/9783110846300.123>
- Kiparsky, Paul, and Carol Kiparsky. 1970. "Fact." In *Progress in Linguistics*, ed. by Mafred Bierwisch, and Karl Erich Heidolph, 143–170. The Hague: Mouton.
- Laca, Brenda. 2010. "Mood in Spanish." In *Mood in the languages of Europe*, ed. by Björn Rothstein, and Rolf Thieroff, 198–220. Amsterdam: John Benjamins. <https://doi.org/10.1075/slcs.120.11lac>
- Mithun, Marianne. 1995. "On the relativity of irrealis." In *Modality in Grammar and Discourse*, ed. by Joan Bybee, and Suzanne Fleischmann, 367–388. Amsterdam: John Benjamins. <https://doi.org/10.1075/tsl.32.16mit>
- Navarro Tomás, Tomás. 1974. *Manual de Entonación Española*. Madrid: Guadarrama.
- Rivero, María Luisa. 1977. "La concepción de los modos en la Gramática de A. Bello y los verbos abstractos en la gramática generativa." In *Estudios de Gramática Generativa del español*, María Luisa Rivero, 69–85. Madrid: Cátedra.

- Rivero, María Luisa, and Arhonto Terzi. 1995. "Imperatives, V-movement and logical mood." *Journal of Linguistics* 31: 301–332. <https://doi.org/10.1017/S0022226700015620>
- Rizzi, Luigi. 1997. "The fine structure of the left periphery." In *Elements of Grammar*, ed. by Liliane Haegeman, 281–337. Dordrecht: Kluwer. https://doi.org/10.1007/978-94-011-5420-8_7
- Roussou, Anna. 2010. "Selecting complementizers." *Lingua* 120: 582–603. <https://doi.org/10.1016/j.lingua.2008.08.006>
- Salvá, Vicente. 1852 [1830]. *Gramática de la lengua castellana*. Madrid: José Ferrer de Orga.
- Sánchez López, Cristina. 2016. "Person features and functional heads. Evidence from an exceptional optative sentence in Ibero-Romance." In *Romance languages and linguistic theory. Selected Papers from the Going Romance, Lisbon 2014*, ed. by Ernestina Carilho, Alexandra Fiéis, Maria Lobo, and Sandra Pereira, 259–278. Amsterdam: John Benjamins.
- Sánchez López, Cristina. 2017. "Optative sentences in Spanish." In *Advances in the Analysis of Spanish Exclamatives*, ed. by Ignacio Bosque, 82–107. Columbus: The Ohio University Press.
- Sánchez López, Cristina. In press. "Optative and Evaluative *que* 'that' sentences in Spanish." In *Insubordination: New perspectives*, ed. by Karin Beijering, Gunther Kaltenböck, and María Sol Sansiñena. Berlin: Mouton De Gruyter.
- Siegel, Laura. 2009. "Mood selection in Romance and Balkan." *Lingua* 119: 1859–1882. <https://doi.org/10.1016/j.lingua.2008.11.006>
- Spaulding, Robert. 1934. "Two Elliptical Subjunctives in Spanish." *Hispania* XVII: 355–360. <https://doi.org/10.2307/332784>
- TLFi *Le Trésor de la langue française [on line]*. <https://atilf.atilf.fr>
- Villalta, Elisabeth. 2007. *Context Dependence in the Interpretation of Questions and Subjunctives*. Ph.D. dissertation, Universität Tübingen.
- Villalta, Elisabeth. 2008. "Mood and Gradability: An Investigation of the Subjunctive Mood in Spanish." *Linguistics and Philosophy* 31: 467–522. <https://doi.org/10.1007/s10988-008-9046-x>

Resumed phrases (are always moved, even with in-island resumption)

Dominique Sportiche

UCLA

On the basis of the properties of resumption in colloquial spoken French, I conclude that resumption of a phrase by a pronominal element always involve run of the mill movement of this phrase, whether the resumptive pronoun is inside an island or not. I propose that such movement can take place in two steps, the first being one of the Left Dislocation options, which feeds a variety of possible second steps (e.g., wh-movement or Clitic Left Dislocation).

Keywords: resumption, movement, island, reconstruction, Dislocation, clitic left dislocation

1. Intro: Resumption

In this short article, I will discuss the syntactic derivation of resumptive constructions. The general schema for resumptive structures as compared with movement constructions is shown in (for illustrative purposes, pseudo-) English below, where X can be an island or not:

- (1) *who ...* [_X ... *you saw a picture of* <gap>...] (* if X is an island)
- (2) *who ...* [_X ... *you saw a picture of him ...*] (X an island or not)

In some languages (e.g. Standard American English), acceptable cases of resumption involve a resumptive pronoun in an island, but this is in no way the general rule cross linguistically. Although I will focus here on wh-constructions (resumptive questions), this extends to some other resumptive constructions in (colloquial spoken) French, e.g. Clitic Left Dislocation (CLLD), restrictive resumptive relatives but not to appositives, and to some other languages too. I will conclude that resumptive structures can always involve run of the mill (wh) movement, regardless of the location of the resumptive pronoun. I will also provide some evidence

that they must, as would independently follow from Ockham's razor/minimalist parsimony.

Space limitations prevent but a brief discussion of the general implications of such conclusions for the analysis of resumption, as well as of one consequence, the background impetus for the present work, namely that resumptive relatives are compatible with the promotion derivation of relative clauses, which Sportiche (2017a) argues is the only derivation available.

1.1 Background

1.1.1 *Reconstruction*

As I will rely on reconstruction phenomena to reach conclusions about the syntactic derivation of resumptive structures, I first want to precisely spell out some of the assumptions I am making about relevant properties of reconstruction.

The most central assumption (cf. Sportiche 2017c, for a survey of motivations) is that (apart from possibly special cases, e.g. pseudo clefts, cf. Sharvit 1999), *reconstruction is a property of movement only*: an item can display displaced interpretation relative to its overt syntactic position only if it has moved from (/to) the position in which it is interpreted.

Regarding scope, I am further assuming, as is standard, that:

- i. a pronoun bound by a quantifying element must be in the scope of this element at the syntax/interpretation interface (LF).
- ii. the *de dicto* reading of some expression relative to some intensional element is licensed only if it is in the scope of this intensional element.

The informative cases are cases in which an expression containing a bound pronoun or an expression read *de dicto* is not in the surface scope of its binder. In such a case, *total reconstruction* of this expression is required: some mechanism must insure that it ends up fully in the scope of its binder. By assumption this mechanism (be it syntactic or semantic reconstruction) crucially references the presence of a trace: reconstructability of an expression for pronominal binding or the possibility of a *de dicto* reading diagnoses the *possibility* for a trace to occur in the scope of the binder.

Next, regarding Condition C effects: If a sentence containing a preposed constituent with a name in it displays a non-coreference effect between this name and a pronoun not c-commanding it, this diagnoses the *necessity* for there to be an *A bar* movement trace lower than the pronoun. If such a trace was not required, there would be some derivation not violating Condition C. Furthermore, if only an A-trace was in the c command domain of the pronoun, we would not

expect a condition C effect as A movement can (under certain circumstances) bleed condition C.

1.1.2 *Analyses of resumption*

Consider a resumptive structure clause. To simplify, assume that there is some antecedent DP located at the periphery of the resumptive structure, RP the resumptive pronoun related to it:

$$DP_k \dots RP_k$$

This RP and the DP must end up coindexed,¹ and other interpretive mechanisms then kick in to derive the appropriate representation.

There are two broad types of analyses regarding how RP is linked to DP.

First this could be achieved via an *Anaphoric Dependency*.

This anaphoric dependency could be a *Binding* relation akin to pronominal binding, as in McCloskey (1990) or Sells (1984): DP binds RP.

This anaphoric dependency could arise via *Agree* as in Rouveret (2008) or Adger and Ramchand (2005): the RP agrees with a peripheral probe and this relation is translated into the needed representation (for concreteness the probe is coindexed with DP).

This anaphoric dependency could be one of *identity under Ellipsis* as in, e.g., Guillot and Malkawi (2007): they take RP to be a definite version of DP, where the NP restriction of the RP's definite article deletes under identity with the NP of the DP yielding a pronoun as in Elbourne (2001).

Secondly, this relation could be established via *Movement*.

In Aoun, Choueiri and Hornstein's (2001) version, the DP originates as *doubling*, and thus coindexed with RP yielding [RP DP] as in "big DP" analyses. The DP subextracts from this big DP and undergoes normal movement.

In other versions, *RP itself moves, covertly* for Demirdache (1991, 1997), *overtly* for Demirdache and Percus (2008, 2011). Movement of RP to the periphery creates a property (a lambda abstract) predicated of the DP: this yields the effect of the desired coindexing between DP and RP.

I will conclude in favor of a movement analysis, but different from either of the two options above: I will argue that movement of the DP (and not of RP) is always involved; and I will conclude that movement is not necessarily from a doubling position.

1. This is good enough for our purposes here and will simplify exposition. To be more precise, in relatives, we would minimally need to have the head NP of the relative coindexed with the implicit NP restriction of the gap or the RP. All arguments about reconstruction will be about this NP.

2. Resumption in islands: Mixed two-step

Cross linguistically resumption is found with resumptive pronouns inside or outside islands. Consider again the DP/RP configuration and let us examine the case in which the RP is separated from DP by an island boundary:

- (3) $DP_k [_{XP} \dots [_{\text{island boundary}} \dots RP_k \dots$

Let us accept for now as is the norm (I will question it later) that islands are barriers for Movement/Agree relations: this excludes an Agree, a “big DP” doubling movement or an RP movement approach to the DP/RP link. But such structures are compatible with a binding approach of RP by DP, or an ellipsis of the RP’s NP under identity with the DP’s NP.²

More generally, it is perhaps now a universal consensus also that, when the RP is inside an island, there is no movement dependency linking DP to a position inside its sister constituent, here XP. The following data shows that this is incorrect. Consider the following (colloquial spoken) French pair of questions:

- (4) a. *Quelle photo de fiançailles_k Jean pense que si son_k auteur vient, on est foutus?*
 ‘What engagement picture John thinks that if its author comes, we are doomed?’
- b. [*Quelle photo de lui_m*]_k *j’ai dit à aucun accusé_m que si son_k auteur vient, on est foutus?*
 ‘[What picture of him_m]_k I told no one_m accused that if its_k author comes, we are doomed?’

In both, there is a resumptive pronoun RP *son/his* inside an island, namely the preposed conditional clause (underlined). In (4a), *photo de fiançailles/engagement picture* can be interpreted *de dicto*: John thinks there is such a picture (among several) but in fact there aren’t any, as I the speaker, know; the picture he is thinking about is in fact a wedding picture. Given our assumptions about reconstruction, we must conclude that in (4a), there is a possible derivation in which the DP *photo de fiançailles/engagement picture* has moved from a position in the scope of the verb *penser/think*. This means that with this interpretation, the derivation must have proceeded as in:

- $DP_k \dots \text{penser} \dots \text{trace}_{DPk} \dots [_{\text{island boundary}} \dots RP_k \dots$

2. A possible exception is covert movement of the RP as in Demirdache (1991, 1997), but covert movement I assume is no longer an available option, given minimalist guidelines.

In (4b), it is possible to interpret the pronoun as bound by the QP, demonstrating the presence of a trace. This means that with this interpretation, the derivation must have proceeded as in:

$$[_{DP_k} \dots \text{pronoun}_{m\dots}] \dots QP_m \dots \text{trace}_{DP_k} \dots [_{\text{island boundary}} \dots RP_k \dots]$$

What these examples show is the existence of a movement derivation (non vacuously) preposing DP from some position P in structures such as (3), if the RP is inside an island: extraction can be from a position non local to the RP:

$$(5) DP_k \dots [_P t_{DP_k}] \dots [_{\text{island boundary}} \dots RP_k \dots]$$

I will call such derivations *mixed two-step* or *MTS* derivations.³ Indeed, the link between P and RP cannot (normally) be a movement relation, while the relation between DP and P must be.

Furthermore, it can be shown that at least in some cases,⁴ DP *must* have moved from a low enough position regardless of where the (NP of the) DP scopes. Indeed, Condition C effects obtain, as they normally do with in *wh*-questions, when the RP is inside an island (underlined):

$$(6) * [\text{Quelle photo de Jean}_{m}]_k \text{ il}_{m} \text{ pense que si tu la}_{k} \text{ brûles, t'es pas un ami}$$

‘[What picture of John_m]_k he_m thinks that if you burn it, you’re not a friend’

This means that MTS derivations with RP inside an island are mandatory. This conclusion regarding the availability and necessity of MTS derivations raises a number of questions that I will address in turn:

- a. In (5), RP is located inside an island. Are MTS derivations (in principle) available and required when RP is not inside an island? I will answer affirmatively.
- b. Is there independent crosslinguistic evidence for such derivations: I will conclude positively by pointing to cases in Irish and Selayarese that display exactly this pattern of extraction.
- c. What can be said about the position P from which extraction can take place? Based on the crosslinguistic evidence for MTS derivations, I will tentatively answer that such positions are (typically clause-) peripheral topics.
- d. What impact does the existence of such derivations have on the the analysis of resumption? I will conclude that if DP is moved when linked back to a gap, as e.g. in relative clauses, DP must also move in RP structures, albeit possibly from a P-like position away from where the RP is.

3. I am only concerned here with cases in which each step is from an A-bar position down.

4. It is easy to show this with *wh*-questions, or Clitic Left Dislocation. There are complications with restrictive relatives, which do not show Condition C effects as straightforwardly.

3. Mixed two-step without islands

To investigate whether MTS derivations are (in principle) available when the RP is not inside an island, let us first place the RP in a movement accessible position⁵ and check whether the right kind of reconstruction effects obtain.⁶

- (7) a. [*Quelle photo de fiançailles*]_k Jean *pense que toi, tu veux* (*la_k*) brûler
 ‘What engagement picture John thinks that you, you want to burn (it)’
 b. [*Quelle photo de lui*]_m]_k j’ai dit à aucun accusé_m que les flics, ils (l’) ont retrouvé
 ‘[What picture of him]_m]_k I told no one_m accused that the cops, they found (it)’

And indeed, they do: in both, the NP (bold) can scope under the *de dicto* operator *penser*: *photo de fiançailles* can scope under *penser*, thus allowing a *de dicto* reading (John is mistaken about the nature of this picture). *Photo de lui* can scope under *aucun accusé* allowing a bound reading for the pronoun. I conclude that both preposed wh-phrases can have moved from some low enough position P.

Just as in the island cases, it can be shown that the wh-phrase **must** have moved from a low enough position. Indeed, Condition C effects obtain, as they normally do with wh-questions (regardless of the presence of the RP):

- (8) [*Quelle photo de Jean*]_m]_k *il_m pense que toi, tu veux* (*la_k*) brûler
 [What picture of John]_m]_k he thinks that you, you want to burn (it)

But this does not suffice to show that MTS derivations are available if RP is not inside an island. Indeed, what could be happening in such cases is homogeneous wh movement steps proceeding from a position P forming e.g. a “big DP” [RP DP] with RP. This is what Aoun et al. (2001) argue is the case in Lebanese Arabic in such configurations. This would fit their schema:

- (9) DP_k ... [CP ... [bigDP RP_k [P t_{DPk}] bigDP]
 where

5. This may not always be feasible. Indeed, it has long been noted (cf. Rouveret 2011, for a recent survey) resumption can (at least appear) to be a Last Resort option, at least in some languages: accordingly, resumption would thus be allowed only if movement is not (cf. Shlonsky 1992). Additional factors are relevant too. In my (colloquial spoken) French, resumption markedly degrades if the RP is in a movement accessible position in the immediate clause under its antecedent, and perhaps even more if the RP is the subject (the latter case recalling the “highest subject constraint” on resumption).

6. Note that despite the presence of the embedded dislocation, the embedded clause is not an island for non resumptive wh-movement. Both sentences are fine without the RP.

- i. DP is subextracted from big DP
- ii. ... does not include island boundaries
- iii. movement proceeds successive cyclically via the CP edge(s)

Consider a movement derivation out of a non island involving several steps:

- (10) $DP_k \dots [{}_P t_{DP_k}] \dots RP_k \dots$
 | --- step2 ---| ---- step 1 ----|

To show that this qualifies as an MTS derivation, it must be shown that two steps (perhaps need not but) can have different properties, e.g. that steps 1 and 2 above, the step from RP to P and the step from P to DP, can have different properties. We also know, because of the island MTS cases above, that there is no principled reason for the P/RP distance to be bounded (since P and RP can be separated by one, or more, island boundaries). In particular, the distance P/DP may be trivial (if P is already at the left periphery of the main clause).

Let us follow Iatridou (1995), who used Parasitic Gaps (PG) licensing to justify MTS derivations (to handle CLLD in Greek out of non-islands).⁷ I will exhibit cases in which step 1 does not license PG but step 2 can. Following Nissenbaum (2000), I take it – for the cases I am looking at – that the crucial properties licensing PGs in VP adjuncts is the presence of a VP peripheral trace of the licensing A bar movement. This allows the VP and the adjunct to type-match for predicate modification composition and to be predicated of this trace:

- (11) $DP_k \dots [{}_{VP} t_k [{}_{VP} \dots [{}_{VP} V t_k]] [{}_{VP-adjunct} e_{PG} [\dots t_{PG}]]]$

Example (12a) is a standard PG licensing configuration used as benchmark. In (12b) (for some reason not as good, possibly because high attachment of the adjunct is less easily accessible) the PG containing constituent is adjoined to the *vouloir/want* clause (as shown by the desired interpretation and the approximate bracketing) and is licensed by wh movement in the *vouloir/want* clause:

- (12) a. *Dis moi ce qu'il veut que tu [comprendes t sans mémoriser pg]*
 Tell me what he wants you to [understand t without (you) memorizing pg]
 b. *?Dis moi ce qu'il [[veut que tu comprendes t] sans même mentionner pg]*
 'Tell me what he [wants you to understand t] without (him) even mentioning pg.'

Adding RP to (12a) yields (13a), which is ill-formed. This means that the A bar dependency with the RP in the most embedded clause does not license PG.

7. Angelopoulos and Sportiche (to appear) argues that MTS derivations (of the relevant sort) are not warranted for Greek CLLD.

Crucially adding an RP to (12b) yields (13b) which is comparable to (12b) (in fact slightly better to my ear, possibly because the alternative with the adjunct attached low as in (12b) is not available). This means that the A-bar dependency in the *vouloir/want* clause does license the PG and is thus different from that with the RP in the most embedded clause:

- (13) a. * *Dis moi ce qu'il veut que tu [le comprennes sans mémoriser pg]*
 'Tell me what he wants you to understand it without (you) memorizing pg'
 b. *?Dis moi ce qu'il [veut que tu le comprennes] sans même mentionner pg]*
 'Tell me what he wants you to understand it without (him) even mentioning pg'

In other words, step 2 licenses PGs, but step 1 does not. This shows the existence, hence the availability of an MTS derivation even with RP not in an island.⁸

Now given that MTS derivations are available, consider the following RP example:

- (14) *Dis moi quelle photo de Jean_k il_{*k} veut que tu [la regardes]*
 'Tell me what picture of John_k he_{*k} wants you look at it.'

We observe a Condition C effect. This means that an A-bar trace of *quelle photo de Jean* (/what picture of John) must be present below *il* (/he). This shows that there is no alternative to a movement derivation (although the movement could either be a big DP derivation, or an MTS derivation from P lower than the pronominal subject).⁹

4. Analytical consequences: Movement and resumption

Assume the theory of reconstruction presented above in Section 1. Assuming that what French shows is general means that resumptive structures must always involve movement, albeit not necessarily from, and sometimes demonstrably not

8. That there are mixed steps is not in question. But as a reviewer rightly points out, exactly how (French) clitics are treated matters to the precise characterization of steps 1 and 2. Space limitations prevent me from discussing this here. I refer the reader to Angelopolous and Sportiche (to appear).

9. Space limitations prevent me from discussing this here but the data suggests that in French, at least, step 1 does trigger Condition C effects. This means Condition C does not provide grounds to distinguish between the two derivations.

from, the RP position.¹⁰ This excludes all non-movement analyses (Binding or Ellipsis) and excludes all movement analyses where the relevant DP itself did not move (RP movement). This is only compatible with Aoun et al.'s (2001) "big DP" subextraction analysis (with P adjacent to RP)¹¹, although such analyses cannot be the only ones available as demonstrated by RP in island constructions as well as by the discussion of RP not in islands in French (and Selayarese or Irish below).

5. Where is P and what is it?

Given the above assumptions and data, the existence of the position P from where movement can be launched in RP structures is not in doubt. Given the crucial role that P plays in resumption, the study of resumptive constructions now becomes in part the study of where this P position can be, what interpretation is associated with it and what relation it bears to the RP.

As I will discuss in the next section, Selayarese and Irish make it plausible that P is (or can be) a clause (or phase¹²) peripheral Topic position, suggesting either that Hanging Topic Dislocation, or, more likely, Clitic Left (or actually Right) Dislocation (see Alexiadou 2006, for a comparative survey of their properties) are involved.¹³ Of course, only a detailed investigation in each language will determine the plausibility of the general approach proposed here and the particular instance of Dislocation used, if any.

10. Small, preliminary survey of speakers of Lebanese Arabic or Hebrew (thanks to Lina Choueiri, Danny Fox, Roni Katzir, Yael Sharvit, and Sarah Ouwayda for their kind help) does suggest that the facts leading to MTS derivations hold there as well.

11. Although it turns out that Aoun et al. (2001) does not provide direct evidence for the big DP analysis because it does not provide direct evidence that the moved DP must reconstruct as low as the RP (rather than lower than some binder).

12. DP peripheral topics are found cross linguistically, and also in French, viz. (i) *t'as vu [Jean_k [l'état où il_k est]]/ did you see [John the state he's in]]*, hence the possibility that extraction is from phase peripheral positions.

13. A Focus position is not a priori ruled out either. Although preposed foci typically leave a gap, unlike preposed Topics which use resumption, RPs in some instances (e.g. weak forms in Lebanese Arabic, Hebrew or Welsh) could in fact be agreement markers and thus qualify as gaps.

6. Independent evidence for MTS derivations

I will now briefly provide independent evidence for the existence of MTS derivations based on patterns found in languages where MTS derivations are more easily detectable. This will also help narrow down what the position P is.

There are two well motivated cases in the literature which display the profile of MTS derivations: Selayarese and Irish.

6.1 Selayarese

Finer's (1997) abstract demonstrates the relevance of Selayarese:

This paper deals with *wh*-questions in Selayarese, discussing two superficially similar constructions that display different properties along the path between the *wh*-operator and the gap. The constructions behave differently with respect to weak crossover, the occurrence of overt complementizers and agreement affixes, and word order. I argue that these differences follow from the hypothesis that the gap in one *wh*-construction is formed by cyclic movement while the other construction involves simple binding between an A' position and a null resumptive element in an A (argument) position. The resumptive A' dependency, however, also displays some surprising movement diagnostics, and so I argue that a *wh*-operator enters the derivation in A' position, binds the A position, and then undergoes movement to its overt position...

Finer (1997) thus illustrates the existence of an MTS derivation. Space limitations prevent me from illustrating in detail some aspects of the syntax of Selayarese relevant to determining the kind of position that P is. I only describe them abstractly. I will not discuss the similarities between the MTS derivation depicted below in (15a) and the successive cyclic (SC) *wh* movement in (15b). They have to do with the fact that the dependency between the intermediate positions P-MTS and P-SC respectively, show all the properties of movement. But the two derivations differ in four respects:

- (15) a. $DP_k \dots [_{P-MTS} t_{DP_k}] \dots [_{CP} \dots V \dots RP_k \dots]$
 b. $DP_k \dots [_{P-SC} t_{DP_k}] \dots [_{CP} \dots V \dots t_k \dots]$
1. The complementizer of CP must be present in (15a) absent in (15b).
 2. Selayarese is V first. But the preverbal position in the embedded clause can be filled by a Focused element in (15a) but not in (15b).
 3. A WCO effect is observed in (15a) but not in (15b).
 4. In both (15a) and (15b), the argument position RP or t is silent, but "definite agreement" is required in (15a), impossible in (15b), and impossible with a clause initial focus as well.

Points 3 and 4 duplicate the distinction found e.g. in Romance languages, where Topicalization (whether CLLD or Hanging Topic) requires resumption (e.g. by a clitic, aka definite agreement) and does not trigger a WCO effect, while focalization precludes clitic doubling and does trigger a WCO effect (see Rizzi 1997). Finer (1997) takes Point 2 to show that (15b) is derived by moving through the embedded preverbal focus position (as Rizzi 1997 does) while (15a) can't be.

Selayarese shows two points: first, it corroborates the existence of MTS derivations with RP not in islands. Second it exemplifies a case in which P is a "clitic doubled" Topic position, as Finer (1997) concludes. Finer (1997) also reports that step 1 of the Selayarese MTS derivation is not movement (reflexives do not reconstruct) and is island creating. This very strongly recalls what Aoun et al. (1998) report of Arabic: one type of CLLD-ed Topics (base generated) have exactly the same two properties, supporting Finer's conclusion that P in Selayarese is indeed such a Topic position.

6.2 Irish

Irish, as analyzed in McCloskey (2002), provides independent support for the existence of MTS derivations with RPs inside or outside islands, as well as information about P. A-bar extractions in Irish display two "standard" patterns A and B exemplified below (from McCloskey 2002):

- (16) a. *an t-ainm a hinnseadh dúinn a bhí ar an áit*
 the name aL was-told to-us aL was _ on the place
 'the name that we were told was on the place'
- b. *an rud a dtug sé orm mionnughadh go gcoinneóchainn*
 the thing aN brought he on-me swear[-fin] go keep [cond]
ceilte é
 concealed it
 'the thing that he made me swear that I would keep hidden'

The key distinction illustrated below has to do with the complementizers, *aL*, *aN* and *go*, on the path from the missing argument to its antecedent, as well as pattern B being limited to DPs.

(17) Standard Patterns

- A: $XP_j [_{CP} aL \dots [_{CP} aL \dots [_{CP} aL \dots t_j \dots]]]$
 B: $DP_j [_{CP} aN \dots [_{CP} go \dots pro_j \dots]]$

McCloskey (1990), for example, documents the properties of these standard patterns as follows:

Pattern A: $XP_k [_{CP} aL \dots [_{CP} aL \dots [_{CP} aL \dots t_k \dots]]]]$

- i. XP binds a gap (trace);
- ii. the XP/trace relation shows all the standard properties of movement to an A' position: island effects, weak crossover, and so on;
- iii. Intermediate complementizers marked *aL* display morphosyntactic evidence typical of successive-cyclic movement.

Pattern B: $DP_k [_{CP} aN \dots [_{CP} go \dots pro_k \dots]]$

- i. DP binds a pronoun (without movement);
- ii. No evidence for a movement dependency DP/pronoun: immunity from islands, absence of weak crossover effects (McCloskey 1990);
- iii. morphosyntactic evidence (in the form of the default complementizer *go*) that intermediate C-positions play no role in establishing the binding relation. Note however the form of the highest complementizer *aN*.

The difference in island sensitivity is illustrated below (from Sells 1984: 200–201):

(18) a. *aL*: island sensitive

**an fear a bpóg me an bhean a phós*
 the man *aL* kissed I the woman *aL* married
 'the man that I kissed the woman that married'

b. *aN*: not island sensitive

an fear a bpóg me an bhean a phós é
 the man *aN* kissed I the woman *aL* married him
 'the man that I kissed the woman that married him'

MTS derivations are illustrated by the existence of mixed chains (pattern #2 in McCloskey 2002, Section 5.2), which constitute a morphosyntactically transparent instantiation of derivations which I claim are found in French, Selayarese and elsewhere (cf. fn. 10). They are found with RP in an island (19a), or not (19b):

(19) a. *aon duine a cheap sé a raibh ruainne tobac aige*
 any person *aL* thought he *aN* was scrap tobacco at-him
 'anyone that he thought had a scrap of tobacco'

b. *an galar a chuala mé ar cailleadh bunadh an oileáin leis*
 the disease *aL* heard I *aN* died people the island [gen] by-it
 'the disease that I heard that the people of the island died of (it)'

Abstractly, they display the following schema:

(20) $DP_k [_{CP} aL \dots [_{CP} aN \dots [_{CP} go \dots pro_k \dots]]$
 $DP_k [_{CP} aL \dots t_k [_{CP} aN \dots [_{CP} go \dots pro_k \dots]]$

Between *aN* and *pro*, there is no evidence of movement. But the presence of *aL* demonstrates that movement has taken place from a position P marked here with t_k .

$$DP_k \leftarrow \text{step 2 (wh-mov.)} \leftarrow [{}_P DP_k] \leftarrow \text{step 1} \leftarrow RP_k$$

Given what we see in pattern B, namely that a DP is found to the immediate left of *aN*: $DP_k [{}_{CP} aN \dots [{}_{CP} go \dots pro_k \dots]]$, the natural analysis is to take step 2 to have occurred from this position. While it is difficult to precisely identify the nature of this position, Rizzi's (1997) left periphery analysis makes it plausible to take it to be a high Topic, preceding some complementizers (such as *aN*).

6.3 Conclusion

MTS derivations of resumptive structures are available in French, whether the RP is inside an island or not. This pattern is duplicated in Irish, and to a lesser extent in Selayarese (where we do not know about the non-island case). For French, because of Condition C effects, MTS derivations of at least some resumptive structures (wh-questions) are mandatory.

7. Brief crosslinguistic comparison

I will now briefly compare what is reported in Lebanese Arabic and what is found in Colloquial Spoken French.

According to Aoun et al. (2001), Lebanese Arabic RPs can be found in islands. No reconstruction is allowed inside the island, but as noted here (cf. fn 10), reconstruction is allowed to (a) position(s) outside of the island suggesting MTS derivations. Or RP can be outside of the island and allow reconstruction (perhaps all the way to RP, cf fn. 11). Furthermore, Aoun and Benmamoun (1998) report the existence of two kinds of CLLD, one with a moved CLLD-ed DP, allowing reconstruction (possibly down to the corresponding RP) and sensitive to islands and one with a base generated CLLD-ed DP not giving rise to reconstruction and not sensitive to islands.

These two sets of facts can be related by taking MTS derivations in Lebanese Arabic to involve CLLD followed by e.g. wh-movement of the CLLD-ed phrase. When the RP is inside an island, the CLLD step must be created by non-movement CLLD, precluding reconstruction into the island. When the RP is not inside the island, the second step of an MTS derivation could either be from a moved CLLD-ed phrase or from a base generated CLLD-ed phrase, which should show no reconstruction below this base generated position.

In French, Guillot and Malkawi (2006) note that reconstruction can take place into islands with resumptive structures:

References

- Adger, David, and Gillian Ramchand. 2005. "Merge and Move: Wh-dependencies Revisited." *Linguistic Inquiry* 36: 161–193. <https://doi.org/10.1162/0024389053710729>
- Alexiadou, Artemis. 2006. "Left Dislocation." In *Syncom: The Blackwell Companion to Syntax*, ed. by Martin Everaert, Henk Van Riemsdijk, Rob Goedemans, and Bart Hollebrandse, 668–700. Malden, MA: Wiley. <https://doi.org/10.1002/9780470996591.ch37>
- Angelopoulos, Nikolaos, and Dominique Sportiche. To appear. "French Dislocations are plain (Scrambling) Movements." In *Romance Languages and Linguistic Theory 2016: Selected papers from Going Romance, Frankfurt 2016*, ed. by Martin Elsig, Ingo Feldhausen, Imme Kuchenbrandt, and Mareike Neuhaus. Amsterdam: John Benjamins.
- Aoun, Joseph, and Elabbas Benmamoun. 1998. "Minimality, reconstruction, and PF movement." *Linguistic Inquiry* 29: 569–597. <https://doi.org/10.1162/002438998553888>
- Aoun, Joseph, Lina Choueiri and Norbert Hornstein. 2001. "Resumption, Movement, and Derivational Economy." *Linguistic Inquiry* 32: 371–403. <https://doi.org/10.1162/002438901750372504>
- Cinque, Guglielmo. 1990. *Types of A-dependencies*. Cambridge, MA: MIT Press.
- Demirdache, Hamida. 1991. "Resumptive Chains in Restrictive Relatives, Appositives and Dislocation Structure." Ph.D. dissertation, MIT.
- Demirdache, Hamida. 1997. "Dislocation, Resumption and Weakest Crossover." In *Materials on Left- Dislocation*, ed. by Elena Anagnostopoulou, Hank van Riemsdijk, and Frans Zwarts, 193–231. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.14.12dem>
- Demirdache, Hamida, and Orin Percus. 2008. "When is a pronoun not a pronoun? The case of resumptives." In *Proceedings of the 38th Annual Meeting of the Northeastern Linguistic Society (NELS)*. Vol. 1, ed. by Anisa Schardl, Martin Walkow, and Muhammad Abdurrahman, 231–245. Amherst, MA: GLSA.
- Demirdache, Hamida, and Orin Percus. 2011. "Resumptives, movement and interpretation." In *Resumptive Pronouns at the Interfaces*, ed. by Alain Rouveret, 367–394. Amsterdam: John Benjamins. <https://doi.org/10.1075/lfab.5.10dem>
- Elbourne, Paul. 2001. "E-type anaphora as NP-deletion." *Natural Language Semantics* 9: 241–288. <https://doi.org/10.1023/A:1014290323028>
- Finer, Daniel. 1997. "Contrasting A-Dependencies in Selayarese." *Natural Language and Linguistic Theory* 15: 677–728. <https://doi.org/10.1023/A:1005801115342>
- Guilliot, Nicolas, and Nouman Malkawi. 2006. "When resumption determines reconstruction." In *Proceedings of WCCFL 25*, ed. by Donald Baumer, David Montero, and Michael Scanlon, 168–176. Somerville, MA: Cascadilla Press.
- Guilliot, Nicolas, and Nouman Malkawi. 2007. "Reconstruction without movement." In *Coreference, Modality and Focus*, ed. by Luis Eguren, and Olga Fernandez Soriano, 113–131. Amsterdam: John Benjamins. <https://doi.org/10.1075/la.111.07gui>
- Guilliot, Nicolas, and Nouman Malkawi. 2012. "On covariation & competition: How resumptives distribute and covary?" Handout from the Resumptive Pronouns Workshop.
- Iatridou, Sabine. 1995. "Clitics and Island Effects." *University of Pennsylvania Working Papers in Linguistics* 2 (1), article 3.
- McCloskey, James. 1990. "Resumptive pronouns, A'-binding and levels of representation in Irish." In *The Syntax of the Modern Celtic Languages*, ed. by Randall Hendrick, 199–248. New York: Academic Press.

- McCloskey, James. 2002. "Resumption, successive cyclicity, and the locality of operations." In *Derivation and Explanation in the Minimalist Program*, ed. by Samuel David Epstein and Daniel Seely, 184–226. Oxford: Blackwell. <https://doi.org/10.1002/9780470755662.ch9>
- Nissenbaum, Jonathan. 2000. "Investigations of Covert Phrase Movement." Ph.D. Dissertation, MIT.
- Rizzi, Luigi. 1997. "The fine structure of the left periphery." In *Elements of Grammar*, ed. by Liliane Haegeman, 281–337. Dordrecht: Kluwer.
https://doi.org/10.1007/978-94-011-5420-8_7
- Rouveret, Alain. 2008. "Phasal agreement and reconstruction." In *Foundational Issues in Linguistic Theory*, ed. by Robert Freidin, Carlos P. Otero, and Maria Luisa Zubizarreta, 167–195. Cambridge, MA: MIT Press.
<https://doi.org/10.7551/mitpress/9780262062787.003.0008>
- Rouveret, Alain. 2011. "Some issues in the theory of resumption: A perspective on early and recent research." In *Resumptive Pronouns at the Interfaces*, ed. by Alain Rouveret, 1–62. Amsterdam: John Benjamins. <https://doi.org/10.1075/lfab.5.01rou>
- Sells, Peter. 1984. "Syntax and semantics of resumptive pronouns." Ph.D. dissertation, University of Massachusetts at Amherst.
- Sharvit, Yael. 1999. "Resumptive pronouns in relative clauses." *Natural Language and Linguistic Theory* 17: 587–612. <https://doi.org/10.1023/A:1006226031821>
- Shlonsky, Ur. 1992. "Resumptive pronouns as last resort." *Linguistic Inquiry* 23: 443–468.
- Sportiche, Dominique. 2016. "Movement, Islands, Phase theory and the ECP" Ms., UCLA
- Sportiche, Dominique. 2017a. "Relative Clauses." Ms., UCLA.
- Sportiche, Dominique. 2017b. "Resumption." Ms., UCLA.
- Sportiche, Dominique. 2017c. "Reconstruction, Binding and Scope." In *Syncom: The Blackwell Companion to Syntax* (2nd edition), ed. by Martin Everaert, Henk Van Riemsdijk, Rob Goedemans, and Bart Hollebrandse. Malden, MA: Wiley.

Timing properties of (Brazilian) Portuguese and (European) Spanish

Irene Vogel,¹ Angeliki Athanasopoulou¹ and Natália Brambatti Guzzo²

¹University of Delaware / ²McGill University

Linguistic rhythmic or timing categories, usually defined in terms of isochrony, remain controversial as a meaningful typology for classifying languages, despite decades of research. Romance languages offer an opportunity to address this question since closely related languages are proposed to be at different ends of the typology. We test two such languages: European Spanish (ES) and Brazilian Portuguese (BP). Instead of investigating isochrony *per se*, however, we examine the interface between timing and prominence properties. Since duration is associated with prominence, we test the hypothesis that syllable-timed languages (ES) do not alter duration to express prominence, while non-syllable-timed languages (BP) do. Comparisons of lexical and sentential prominence effects on duration support our hypothesis, confirming the proposed distinction between the rhythmic classes of the two languages.

Keywords: stress-timing, syllable-timing, Brazilian Portuguese, European Spanish, duration, stress, focus, isochrony, rhythm

1. Introduction

The terms syllable- and stress-timing were introduced by Pike (1945) to capture an intuitive rhythmic distinction between languages, establishing a typology characterized by two types of isochrony as the ends of a spectrum. The literal application of this distinction between equal syllable durations and equal inter-stress intervals, however, has not been supported by measurements of speech across languages and remains controversial. Romance languages offer a particularly rich research opportunity since closely related languages are often claimed to be at different ends of the timing or associated rhythmic typology. In some cases, even the same language may appear in different categories (Frota and Prieto 2015). In this paper,

we analyze European Spanish, usually considered “syllable-timed” and Brazilian Portuguese, characterized as “mixed-/stress-timed.” Given the well-documented failure of previous research to identify consistent durational differences, while still recognizing the basic perception of distinct linguistic rhythmic patterns, we propose a different approach to the rhythmic typology: instead of investigating isochrony *per se*, we examine the interface between timing and the manifestation of prominence (i.e., lexical stress and sentential focus). Specifically, we propose that the percept of rhythm must abstract away from extraneous types of variability, and rely only on those duration patterns that provide crucial linguistic information, in particular, the location of stress and focus prominences. Moreover, we incorporate in our assessment of rhythm the fact that speakers / listeners¹ tend to be more sensitive to relative prosodic properties than to absolute values, in this case vowel durations. We thus conduct two types of analyses (i.e., Binary Logistic Regression Analyses and calculations of duration ratios) that allow us to examine the *function* of duration in a language’s prosodic system, as opposed to absolute duration differences.

In the following sections, we first present general considerations regarding syllable- and stress-timing, and the application of this rhythmic classification to European Spanish and Brazilian Portuguese (Section 2). We then present our experimental investigation of the timing properties of these languages: methodology (Section 3) and results (Section 4). In Section 5, we discuss our findings as well as our proposal for the analysis of the rhythmic typology in terms of the function of duration in the expression of prominence. General conclusions follow in Section 6.

2. Rhythmic typology and the classification of European Spanish and Brazilian Portuguese

2.1 Syllable- and stress-timing

Isochrony, the equal timing of different units, has been both proposed and challenged as a system of classification or typology of languages based on their rhythmic characteristics (Pike 1945, Dauer 1983, 1987, Arvaniti 2009). Ideally, in syllable-timed languages, the durations of all syllables are the same, regardless of the locations of stresses, while in stress-timed languages, the durations of the inter-stress intervals are the same, regardless of the number of syllables between the stresses, as represented in Figure 1.

1. Henceforth, we just use the term “speaker” for simplicity.

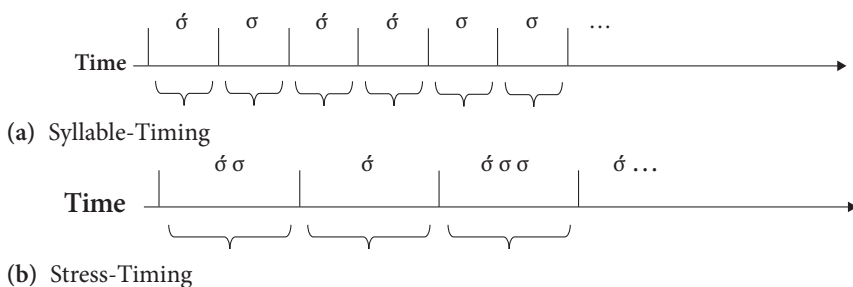


Figure 1. Schematic representation of (a) Syllable-Timing and (b) Stress-Timing.

It should be noted that “timing” refers to a measurable phenomenon, duration in milliseconds, while “rhythm” refers to a perceptual, and / or abstract representational phenomenon. These are thus fundamentally different properties; however, the terms are often used interchangeably on the assumption that the physical, measurable, durations underlie the perception of the associated rhythmic categories. Unless otherwise indicated, the same practice is followed here.

Despite intuitions that languages exhibit one or the other type of rhythm (e.g., Lehiste 1973, 1979, Ramus and Mehler 1999, Ramus, Dupoux and Mehler 2003), measurements of the relevant durations have failed to provide consistent support for such intuitions (among others, Dauer 1983, 1987, Ramus, Nespore and Mehler 1999, Low, Grabe and Nolan 2000, Grabe and Low 2002, Dellwo 2006, Arvaniti 2009). This is not particularly surprising since duration is subject to numerous types of influence, including the inherent durations of segments, syllable structures, the positions of segments in syllables, words and phrases, as well as speaker differences and considerations such as speech rate and style. Moreover, different methodologies and statistical analyses may yield results that are not comparable (e.g., Dauer 1983, Grabe and Low 2002, Ramus et al. 1999, Frota and Vigário 2001).² The fundamental question then, is what aspects of the highly variable speech signal are being recognized and serving as the basis of the impression of two types of linguistic rhythmic patterns – if it is not duration.

As an alternative to duration measurements, Arvaniti (2009) suggests that rhythm is a reflection of more abstract properties such as the grouping and patterning of prominences and metrical structure (e.g., Hayes 1995), although no specifics are provided as to how this approach would be implemented. Prieto et al. (2012) propose that while duration is relevant in rhythmic classifications, it must be viewed in relation to the prosodic structures of a language, specifically, the prosodic heads and boundaries. The present research advances a somewhat similar

2. See Arvaniti (2009), Prieto et al. (2012) and http://www.lfsag.unito.it/ritmo/index_en.html for recent overviews of different methodologies and outcomes.

proposal in that certain aspects of prosodic structure are crucial to the percept of linguistic rhythm.

2.2 Rhythmic classification of European Spanish and Brazilian Portuguese

In previous research, European Spanish (ES) has usually been characterized as syllable-timed (e.g., Pike 1945, Delattre 1965, 1966, Carrió i Font and Ríos Mestre 1991, Ramus et al. 1999, Prieto et al. 2012), although in several cases, stress-timing has been proposed (e.g., Navarro Tomás 1916, Hoequist 1983a,b). Brazilian Portuguese (BP), however, is typically considered stress-timed (Major 1985) or mixed-timed (e.g., Parkinson 1988, Barbosa 2000, Frota and Vigário 2001, Massini-Cagliari 1992), reflecting the fact that it does not exhibit extreme vowel reduction or deletion as in prototypically stress-timed languages such as English and European Portuguese (e.g., Parkinson 1988, Ramus et al. 1999).

In their prosodic structure approach, Prieto et al. (2012) have also concluded that ES is syllable-timed, since it did not show much increase in duration either in (sentence-) accented stressed syllables or at phrasal boundaries, compared to stress-timed English. Based on the same criteria, Catalan, too, was found to be syllable-timed, although it is sometimes considered to have a “mixed” rhythm (e.g., Prieto et al. 2012). BP was not investigated, and while it might be tempting to extend the Catalan conclusions to BP, since it is also sometimes claimed to have a “mixed” rhythm, this is not possible, as it is unclear exactly what “mixed” means.

3. Experimental investigation of rhythm in European Spanish and Brazilian Portuguese

Since the absolute durations of syllables and inter-stress intervals are too variable to be reliable as the basis of a language’s rhythmic classification, we advance an alternate approach which allows us to abstract away from such variation. Specifically, we test the general hypothesis that the role of duration in signaling prosodic prominence (i.e., stress and focus) is what underlies the percept of linguistic rhythm. That is, in syllable-timed languages such as ES, duration is *not* altered as a means of manifesting stress and focus since, by definition, syllable durations must be stable. By contrast, in so-called stress-timed languages such as BP, duration *is* altered, specifically increased, in the manifestation of prominence. Since the rhythm in such languages does not actually correspond to the position of stresses, we henceforth refer to it, along with the somewhat vague category of mixed-timed languages, with the more neutral label “non-syllable-timed.” We thus formulate the specific hypotheses in (1) and (2).

- (1) Hypothesis 1: In BP but *not* in ES, duration is the main property in the prosodic manifestation of stress.
- (2) Hypothesis 2: In BP but *not* in ES, duration is the main property in the prosodic manifestation of focus.

Since speakers are not particularly sensitive to absolute duration, and more generally, perceive prominence as a relative rather than an absolute phenomenon, we also consider the rhythmic classes in terms of relative durations. We thus test two additional specific hypotheses, (3) and (4).

- (3) Hypothesis 3: In ES, the durations of stressed and unstressed vowels are similar, resulting in a ratio close to 1; in BP, the ratio is substantially greater than 1.
- (4) Hypothesis 4: In ES, the durations of focused and non-focused vowels are similar, resulting in a ratio close to 1; in BP, the ratio is substantially greater than 1.

3.1 Methodology

The data analyzed in this study come from larger corpora collected in the context of a cross-linguistic investigation of the acoustic properties of stress and focus. The experimental design and analyses thus conform to the specifications applied to all of the languages in order to permit comparison of the results (Vogel, Athanasopoulou, and Pincus 2016).

3.2 Participants

Ten speakers of both BP (South-East variety; 5 F) and ES (standard northern/central variety; 5 F) were recorded.³ All were university students or recent graduates, 18 to 30 years of age.

3.3 Procedure

The participants, tested individually by a native speaker, saw dialogues containing the stimuli on PowerPoint slides, and read both the question and answer of each dialogue, speaking in a natural way. Between dialogue slides, the participants named pictures of objects shown on filler slides. Their speech was recorded using an external, head-mounted, microphone.

3. One ES speaker's recording was discarded for technical reasons.

3.4 Stimuli

The stimuli for both languages are real 3-syllable words with penultimate stress. To avoid possible confounds due to syllable structure and intervening elements between syllables, the words consist of CV syllables to the extent possible; the target vowels, /i, u, a/, always appear in open syllables. There are 10 words for each target vowel in the stressed and unstressed syllables; the third syllable is not examined to avoid potential confounds with word-final effects. Examples are given in Table 1.

Table 1. Examples of target words with /a/ vowel in each language

	ES	BP
unstressed (Syll 1)	<i>caseta</i> [ka'seta] 'hut'	<i>cachorro</i> [ka'foxo] 'dog'
stressed (Syll 2)	<i>butaca</i> [bu'taka] 'arm chair'	<i>pedaço</i> [pe'dasu] 'piece'

To permit examination of stress without the confound of focus, and to assess the effect of focus, the stimuli were elicited in two focus conditions, yielding 120 items per speaker. The short dialogues used to prime the (narrow) focus and non-focus readings⁴ are illustrated for BP in Figure 2.

3.5 Analyses

Vowel Duration, (mean) F0, Δ F0 (change from beginning to end of vowel), Intensity and Vowel Centralization were measured with Praat.⁵ All measurements were normalized for speaker and vowel differences using z-scores (after log transformation of the duration and F0 measurements), so they could be pooled.

Binary Logistic Regression Analyses (BLRAs) were first conducted to statistically assess the distinguishability (i.e., classification rates) of stressed vs. unstressed vowels and focused vs. non-focused vowels using all of the acoustic properties. Differently from other widely used tests of significance, the BLRAs also provide the percentage of data correctly classified into the relevant categories (i.e., stressed vs. unstressed, focused vs. non-focused). Follow-up BLRAs were subsequently conducted with each property found to be significant ($p < .005$) in the overall BLRAs to assess their individual roles in cueing stress and focus when used as the sole classifier.

4. The target in the non-focus condition appears before a *focused* word to avoid possible post-focal compression.

5. To avoid possible effects of the onset consonant, for all F0 measurements, the first point considered is 10 ms into the vowel.



a. Focus Condition	
 <p>O que a Maria disse de tarde?</p> <p>A Maria disse “cachorro” de tarde.</p>	<p>–What did Maria say in the afternoon?</p> <p>–Maria said “cachorro” in the afternoon.</p>
b. Non-Focus Condition	
 <p>A Maria disse “cachorro” de noite?</p> <p>Não, a Maria disse “cachorro” de tarde, não de noite.</p>	<p>–Did Maria say “cachorro” at night?</p> <p>–No, Maria said “cachorro” in the afternoon, not at night.</p>

Figure 2. Brazilian Portuguese Dialogues, with target *cachorro* ‘dog’ in each focus condition. The prominent word is bolded.

Additional descriptive analyses assessed the acoustic measurements themselves. Based on the average z-score (all speakers and items) for each category and acoustic property, we calculated pseudo raw scores using the mean and standard deviation values of an arbitrarily selected speaker of each language. Since the resulting values are derived from the measurements of all the speakers of the language, they represent the general speech patterns of the language, not those of any one speaker.

The mean durations of each stress and focus category were then used to determine the Stress Ratio (SR) and Focus Ratio (FR), as in (5) and (6). These measures allow us to examine relative changes in duration due to stress and focus, thus abstracting away from much of the variability in absolute duration measurements.

$$(5) \text{ Stress Ratio (SR)} = \frac{\text{Duration of Stressed Vowel}}{\text{Duration of Unstressed Vowel}}$$

$$(6) \text{ Focus Ratio (FR)} = \frac{\text{Duration of Focused Vowel}}{\text{Duration of Non-Focused Vowel}}$$

4. Results

4.1 Binary logistic regression analysis (BLRA) results

Table 2 presents the BLRA rates of classification of Stressed vs. Unstressed vowels in ES and BP, first for all of the acoustic properties combined (Overall classification), then for the top two statistically significant properties when tested as the sole classifier (Individual classification). The crucial condition here is the Non-Focus condition (unshaded), since this shows the stress properties without the possible confound of focus properties. (Dur = duration, F0 = mean F0, Δ F0 = F0 change).

Table 2. Classification of stressed vs. unstressed vowels

	Stressed vs. Unstressed			
	Non-Focused		Focused	
	Overall	Individual	Overall	Individual
ES	86%	F0 (82%), Δ F0 (76%)	89%	F0 (86%), Δ F0 (81%)
BP	73%	Dur (68%), Δ F0 (62%)	88%	Dur (84%), F0 (81%)

Both languages show a clear distinction between the Stressed and Unstressed vowels, although the overall classification rate is considerably higher in ES (86%) than in BP (73%). The languages differ strikingly, however, in which property is primarily responsible for the distinction: F0 in ES, but Duration in BP.

Table 3 presents the BLRA results for the classification of Focused vs. Non-Focused vowels. The crucial findings (unshaded) are those that combine the properties of stress and focus, since the stressed vowel is the main locus of focus enhancement (e.g., Ladd 1996).

Table 3. Classifications of focused vs. non-focused vowels

	Focused vs. Non-Focused			
	Unstressed		Stressed	
	Overall	Individual	Overall	Individual
ES	67%	Dur (60%), F0 (60%)	78%	Dur (73%), Int (65%)
BP	89%	F0 (85%), Δ F0 (67%)	83%	Dur (79%), F0 (67%)

Both languages again exhibit strong overall classification rates (ES = 78%; BP = 83%), and in this case, they make use of the same main classifying property, Duration.

4.2 Duration

Since the BLRAs provide information about statistical significance, but not the actual acoustic qualities of the vowels, we next consider the (normalized) durations of the stress and focus categories (Figure 3).⁶

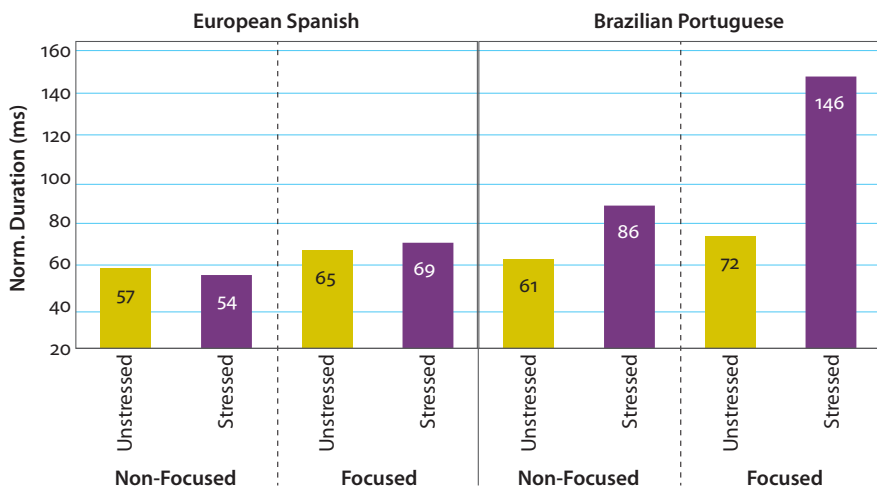


Figure 3. Normalized Durations (ms).

In the crucial, non-focus, condition for stress, the durations of the stressed (54 ms) and unstressed vowels (57 ms) in ES are essentially the same, as expected from the BLRA classification, where duration was not significant. In BP, however, the difference was significant, reflecting the increase from 61 to 86 ms (41%).

In the relevant position for examining focus, the stressed syllable, consistent with the BLRAs, there are significant differences in both languages, with duration increased under focus. In BP, the difference between the non-focused (86 ms) and focused (146 ms) vowels shows a substantial increase of 60 ms (70%); however, in ES, the increase is much smaller (non-focused = 54 ms; focused = 69 ms): 15 ms (28%). While this result is statistically significant in the BLRA, it is not clear whether it is consistently perceptible.

6. While F0 is also an important property in the BLRAs, we focus here on duration in examining rhythm and timing.

4.3 Duration ratios

As noted above, speakers must abstract away from most variation in vowel duration. Attending to relative rather than absolute duration changes offers such an option, since ratios of durations tend to exhibit more overall consistency than absolute durations (e.g., Lehiste 1973, 1979). The Stress and Focus Ratios are shown in Table 4; the crucial positions (as identified above) are unshaded.

Table 4. Stress (SR) and Focus Ratios (FR)

Stress Ratio (SR)	Non-Focus Condition	Focus Condition
ES	54/57 = 0.95	69/65 = 1.06
BP	86/61 = 1.41	146/72 = 2.03
Focus Ratio (FR)	Unstressed Position	Stressed Position
ES	65/57 = 1.14	69/54 = 1.28
BP	72/61 = 1.18	146/86 = 1.70

Both timing ratios are close to 1 in ES (SR = 0.95; FR = 1.28), while the ratios in BP are considerably higher (SR = 1.41; FR = 1.70). Even in the condition with the focus confound, the SR in ES (1.06) is close to 1, compared to 2.03 in BP. Moreover, in ES, focus similarly enhances the duration of both stressed and unstressed syllables, so the FR remains close to 1 (stressed = 1.28; unstressed = 1.14). In BP, focus differentiates between stressed and unstressed syllables, barely enhancing the latter (FR = 1.18).

5. Discussion

The different analyses presented above provide evidence that ES and BP are crucially different with regard to the role of duration in the manifestation of prominence. Indeed, the findings support the general hypothesis that duration is not fundamentally altered by stress and focus in a syllable-timed language, while it is substantially enhanced in a non-syllable-timed language.

5.1 Manifestation of prominence in European Spanish and Brazilian Portuguese

The BLRAs showed that duration is a main property in the prosodic manifestation of stress in BP but not in ES (Table 2), strongly supporting Hypothesis 1. Duration was also found, however, to be the main property in the prosodic manifestation of

focus in both BP and ES, not just BP (Table 3), at first glance seeming to conflict with Hypothesis 2. Closer examination of the actual durations, however, revealed that the increase in ES is not very large (15 ms), and possibly not perceptible. Thus, we cannot reject Hypothesis 2.

It should be noted that the finding that ES does not use duration as a significant cue for stress, preferring F0, is also reported elsewhere (e.g., Llisterri et al. 2003, Vogel et al. 2016), as is the finding that BP does use duration (e.g., Major 1985, Massini-Cagliari 1992, Moraes 1998, Cantoni 2013). It has also been suggested, however, that ES uses duration to cue stress (e.g., Ortega-Llebaria 2006, Ortega-Llebaria and Prieto 2011), conflicting with the view of ES as a syllable-timed language, although closer examination reveals that the findings in question may reflect certain methodological and analytical choices. (See discussion in Vogel et al. 2016.)

The relative duration patterns provide support for Hypothesis 3: in ES, the durations of stressed and unstressed vowels are similar, resulting in a ratio close to 1; in BP, the ratio is substantially greater than 1. Hypothesis 4 was also confirmed, since in ES the durations of focused and non-focused vowels are similar, resulting again in a ratio close to 1, while in BP, the durations are quite different, resulting in a ratio considerably greater than 1 (see SRs and FRs in Table 4).

In sum, in ES, the overall duration patterns are not substantially altered by either stress or focus, while in BP they are. Thus, the percept of syllable-timed rhythm (as in ES) derives from the *stability* (not equality) of the vowel durations, while the percept of non-syllable-timed rhythm (as in BP) derives from the alterability of the vowel durations.

5.2 Measurements of relative duration

The present proposal abstracts away from much of the variation in absolute duration measurements, using instead relative duration patterns, the Stress and Focus Ratios. These ratios identify particular duration comparisons as crucial for rhythm, specifically, those related to the manifestation of prominence, discussed further in Section 5.3.

Although a number of previous investigations of rhythm have also considered various types of relative duration properties, they differ from the present proposal, examining durations independently of their prominence-marking function. For example, Ramus et al. (1999) considered the percentages of vowel and consonant segments (%V and %C) in speech samples, as well as the standard deviations of vocalic and inter-vocalic strings. Since these measures were found to be quite sensitive to influences such as speech rate, more complex calculations were developed to attempt to normalize for these additional factors (Dellwo 2006). Comparison

of the results of such analyses shows, however, that depending on which calculation is used, languages may be classified differently with respect to their rhythmic properties. Some languages seem to be better classified by one method and others by another, depending on which outcome coincides with one's assumptions about the rhythmic properties of specific languages.

This subjective approach to rhythmic typology is not surprising since there is thus far no independent characterization of rhythmic classes to replace the problematic concept of isochrony. In fact, the various calculations essentially represent particular phonotactic properties that have been associated with syllable- and stress-timed languages (Dauer 1987). For example, %V and %C reflect syllable complexity and vowel variability measures may reflect the presence of vowel reduction. While such distributional patterns may commonly align with different rhythmic classes, they do not themselves constitute the basis for the classes.

Further consideration of the various relative duration calculations shows, moreover, that they all (except Prieto et al. 2012) rely on the surface duration measurements available in a given string of speech, making the assumption that this is what is relevant for determining rhythmic classes. They, furthermore, typically attribute the same status to elements in different types of syllables, and in different positions in words, phrases and utterances, pooling their durations for the statistical manipulations and analyses. Since any approach that relies on surface durations is sensitive to the nature of the speech sample, it cannot provide a suitable basis for a rhythm-based typology, which requires independent and consistent definitions of the relevant classes.

5.3 Rhythm as duration manipulation

Although our data also consist of surface duration measurements, we propose that it is not simply the duration measurements that must be considered in determining rhythmic classes, but the importance of the durations in the broader context of the prosodic structure of a language. Specifically, the Duration Ratios incorporate information about the role of duration in signaling specific (prosodic) linguistic functions (i.e., stress and focus prominence). Moreover, since they provide relative rather than absolute duration values, they not only avoid much extraneous variability in surface durations, but also accord better with the concept of prominence as a relative phenomenon.

Prieto et al. (2012) also associate rhythm with prosodic structure, proposing that prosodic heads and boundaries are lengthened more in stress-timed than syllable-timed languages. While this approach permits the comparison of languages in terms of how much lengthening they exhibit, it does not provide a metric by

which the rhythmic properties of a language may be identified independently of particular comparator languages.

While we are in agreement with Prieto et al. that duration is substantially greater in stress and focus conditions (Prieto et al.'s stressed unaccented and stressed accented syllables) in so-called stress-timed (here non-syllable-timed) languages, we propose a stronger, causal, relationship between prosody and duration. Specifically, in non-syllable-timed languages, duration is intentionally⁷ employed to signal stress and focus prominence. Thus, what characterizes their rhythm is not isochrony (i.e., equal durations) of inter-stress intervals, but rather the systematic manipulation of duration for the purpose of cueing stress and focus. By contrast, syllable-timed rhythm is characterized by the absence of duration enhancement as a means of cueing prominence, the result being an overall syllable stability, but not necessarily equal durations, or syllable isochrony.

Furthermore, the relational approach to the function of duration provided by the Stress and Focus Ratios offers the possibility of a more nuanced approach to rhythmic typology. Since ratios are in principle infinite, instead of providing only a binary choice of rhythmic classes, syllable- or non-syllable-/stressed-timed, they provide a principled, quantitative, means of characterizing languages, including those that may not solidly fit into the basic categories.

Specifically, with regard to ES and BP, we can plot and compare the SR and FR values (Section 4.3) along the continuum shown in Figure 4. A ratio of 1 represents the ideal for syllable-timing: no change of duration for stress or for focus. No ideal is presently established for the opposite end of the spectrum, since it is anticipated that the value will ultimately be based on a “classic” stress-timed language such as English.

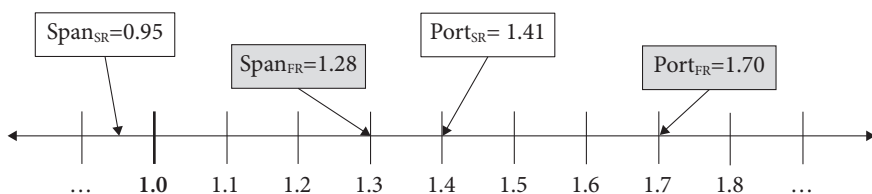


Figure 4. Syllable-timing continuum: Stress (SR) and Focus Ratios (FR) for ES (Span) and BP (Port).

Both the Stress and Focus Ratios are considerably closer to 1 in ES than BP. Moreover, even the higher FR value in ES is lower than the shorter SR value in

7. Note that “intentional” does not mean conscious, but only that duration increases are carried out with a specific grammatical purpose.

BP. Thus, both ratios provide support for a clear rhythmic distinction between the two languages.

Although Figure 4 encodes two continua, which could be displayed separately for the SR and FR, the combined representation allows us to more directly compare the two types of prominence. Moreover, it provides a means of characterizing languages with mixed or intermediate rhythmic patterns: those with SRs and / or FRs around the middle of the continuum, or strongly divergent from each other. Finally, it should be noted that since a ratio calculation will always yield some value, the statistical analyses of the BLRAs provide crucial information regarding the significance of duration in the manifestation of prominence in a language, as well as its relative role in comparison with other prominence properties.

6. Conclusions

Despite the fact that duration measurements have not reliably yielded support for isochrony, the percept of distinct rhythmic or timing categories (syllable- vs. stress-timing) has persisted since its introduction by Pike (1945). We have proposed that this percept rests on linguistic properties that abstract away from numerous types of variability that affect the absolute durations of syllables and inter-stress intervals in speech. Specifically, the classification of languages as syllable- or non-syllable-timed is determined by the use of duration for the purpose of expressing prosodic prominence relations in a language. Syllable-timed languages (e.g., European Spanish) are perceived as such not because the duration of all syllables is the same, but because their vowel durations resist alteration as a means of cueing word and sentence level prominence (stress and focus). These languages contrast with non-syllable-timed languages (e.g., Brazilian Portuguese), which do not resist the manipulation of duration in manifesting prominence, but rather enhance it precisely to cue stress and focus.

Moreover, since we view rhythm in relation to prominence, and prominence is not an absolute property, we developed Stress and Focus Ratios to examine the relative roles of duration in ES and BP. The ratios are displayed on a continuum and thus provide a means of characterizing languages, both independently and in relation to other languages. This innovation also offers insight into languages described as exhibiting “mixed” or intermediate rhythmic properties. Specifically, such languages have ratios in the mid-range of the continuum, or substantially different ratios for stress and focus. Ultimately, of course, the validity of the proposed alternative approach to the rhythmic classification of languages must be tested with other Romance languages, as well as with unrelated languages.

References

- Arvaniti, Amalia. 2009. "Rhythm, Timing and the Timing of Rhythm." *Phonetica* 66 (1–2): 46–63. <https://doi.org/10.1159/000208930>
- Barbosa, Plínio. 2000. "'Syllable-timing in Brazilian Portuguese': uma crítica a Roy Major ['Syllable-timing in Brazilian Portuguese': a critic to Roy Major]." *DELTA* 16 (2): 369–402. <https://doi.org/10.1590/S0102-44502000000200006>
- Cantoni, Maria Mendes. 2013. "O acento no Português Brasileiro: uma abordagem experimental [Stress in Brazilian Portuguese: An experimental approach]." Ph.D. dissertation, Universidade Federal de Minas Gerais.
- Carrió i Font, Mar, and Antonio Ríos Mestre. 1991. "Compensatory Shortening in Spanish Spontaneous Speech." In *Proceedings of the Workshop on Phonetics and Phonology of Speaking Styles*, ed. by Joaquim Llisterra, and Delors Poch, 1–5. Barcelona: ESCA.
- Dauer, Rebecca. 1983. "Stress-Timing and Syllable-Timing Reanalyzed." *Journal of Phonetics* 11: 51–62.
- Dauer, Rebecca. 1987. "Phonetic and Phonological Components of Language Rhythm." In *Proceedings of the 11th International Congress of Phonetic Sciences*, 447–450. Tallinn: Academy of Sciences of the Estonian S.S.R.
- Delattre, Pierre. 1965. *Comparing the Phonetic Features of English, German, Spanish and French*. Heidelberg: Julius Groos.
- Delattre, Pierre. 1966. *Studies in French and Comparative Phonetics*. The Hague: Mouton.
- Dellwo, Volker. 2006. "Rhythm and Speech Rate: A Variation Coefficient for ΔC ." In *Proceedings of the 38th Linguistic Colloquium, Budapest*, ed. by Pawel Karnowski, and Imre Sziget, 231–241. Frankfurt: Peter Lang.
- Frota, Sónia, and Pilar Prieto (eds.). 2015. *Intonation in Romance*. New York: Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780199685332.001.0001>
- Frota, Sónia, and Marina Vigário. 2001. "On the Correlates of Rhythmic Distinctions: The European/Brazilian Portuguese Case." *Probus* 13 (2): 247–275. <https://doi.org/10.1515/prbs.2001.005>
- Grabe, Esther, and Ee Ling Low. 2002. "Durational Variability in Speech and the Rhythm Class Hypothesis." *Papers in Laboratory Phonology* 7: 515–546.
- Hayes, Bruce. 1995. *Metrical Stress Theory: Principles and Case Studies*. Chicago: University of Chicago Press.
- Hoequist, Charles. 1983a. "Durational Correlates of Linguistic Rhythm Categories." *Phonetica* 40 (1): 19–31. <https://doi.org/10.1159/000261679>
- Hoequist, Charles. 1983b. "Syllable Duration in Stress-, Syllable- and Mora-Timed Languages." *Phonetica* 40 (3): 203–237. <https://doi.org/10.1159/000261692>
- Ladd, Robert. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- Lehiste, Ilse. 1973. "Rhythmic Units and Syntactic Units in Production and Perception." *Journal of the Acoustical Society of America* 54 (5): 1228–1234. <https://doi.org/10.1121/1.1914379>
- Lehiste, Ilse. 1979. "The Perception of Duration within Sequences of Four Intervals." *Journal of Phonetics* 7 (4): 313–316.
- Llisterra, Joaquim, María Machuca, Carme de la Mota, Montserrat Riera, and Antonio Ríos. 2003. "The Perception of Lexical Stress in Spanish." In *Proceedings of the 15th International Congress of Phonetic Sciences*, ed. by Maria-Josep Solé, Daniel Recasens, and Joaquín Romero, 2023–2026. Barcelona: Causal Productions.

- Low, Ling Ee, Esther Grabe, and Francis Nolan. 2000. "Quantitative Characterizations of Speech Rhythm: Syllable-Timing in Singapore English." *Language and Speech* 43 (4): 377–401. <https://doi.org/10.1177/00238309000430040301>
- Major, Roy. 1985. "Stress and Rhythm in Brazilian Portuguese." *Language* 61 (2): 259–282. <https://doi.org/10.2307/414145>
- Massini-Cagliari, Gladis. 1992. *Acento e ritmo* [Stress and rhythm]. São Paulo: Contexto.
- Moraes, João Antônio de. 1998. "Intonation in Brazilian Portuguese." In *Intonation systems: A Survey of Twenty Languages*, by Daniel Hirst, and Albert Di Cristo, 179–194. Cambridge: Cambridge University Press.
- Navarro Tomás, Tomás. 1916. "Cantidad de las vocales acentuadas [The quantity of stressed vowels]." *Revista de Filología Española* III: 387–408.
- Ortega-Llebaria, Marta. 2006. "Phonetic Cues to Stress and Accent in Spanish." In *Selected Proceedings of the 2nd Conference on Laboratory Approaches to Spanish Phonetics and Phonology*, ed. by Manuel Díaz-Campos, 104–118. Somerville, MA: Cascadilla Press.
- Ortega-Llebaria, Marta, and Pilar Prieto. 2011. "Acoustic Correlates of Stress in Central Catalan and Castilian Spanish." *Language and Speech* 54 (1): 1–25. <https://doi.org/10.1177/0023830910388014>.
- Parkinson, Stephen. 1988. "Portuguese." In *The Romance Languages*, ed. by Martin Harris, and Nigel Vincent, 131–169. New York: Routledge.
- Pike, Kenneth. 1972 [1945]. "The Intonation of American English." In *Intonation*, by Dwight Le Merton Bolinger, 53–83. Harmondsworth: Penguin.
- Prieto, Pilar, Maria del Mar Vanrell, Lluïsa Astruc, Elinor Payne, and Brechtje Post. 2012. "Phonotactic and Phrasal Properties of Speech Rhythm. Evidence from Catalan, English, and Spanish." *Speech Communication* 54 (6): 681–702. <https://doi.org/10.1016/j.specom.2011.12.001>
- Ramus, Franck, Emmanuel Dupoux, and Jacques Mehler. 2003. "The Psychological Reality of Rhythm Classes: Perceptual Studies." In *Proceedings of the 15th International Congress of Phonetic Sciences*, ed. by Maria-Josep Solé, Daniel Recasens, and Joaquín Romero, 337–342. Barcelona: Causal Productions.
- Ramus, Franck, and Jacques Mehler. 1999. "Language Identification with Suprasegmental Cues: A Study Based on Speech Resynthesis." *The Journal of the Acoustical Society of America* 105 (1): 512–521. <https://doi.org/10.1121/1.424522>
- Ramus, Franck, Marina Nespó, and Jacques Mehler. 1999. "Correlates of Linguistic Rhythm in the Speech Signal." *Cognition* 73 (3): 265–292. [https://doi.org/10.1016/S0010-0277\(99\)00058-X](https://doi.org/10.1016/S0010-0277(99)00058-X)
- Vogel, Irene, Angeliki Athanasopoulou, and Nadya Pincus. 2016. "Prominence, Contrast and the Functional Load Hypothesis: An Acoustic Investigation." In *Dimensions of Phonological Stress*, ed. by Jeffrey Heinz, Rob Goedemans, and Harry van der Hulst, 123–167. Cambridge: Cambridge University Press.

Language index

A

Aragonese 245, 248, 251, 252,
254, 255, 256, 260, 262

C

Catalan 1, 18, 39, 50, 108n11,
122n37, 130, 159n22, 170n48,
201, 202, 204, 206, 209,
210, 211, 212, 232, 234, 235,
239, 240, 241, 243, 244, 245,
246, 251, 252, 253, 254, 255,
256, 257, 258, 259, 262, 306,
328, 340

Chinese 30, 86, 100, 151

Czech 216n2, 228

D

Deliano 55, 56, 57, 61, 62, 63,
64, 65, 66

Dutch 39, 50, 181, 284, 291

E

English 11, 21, 22, 23n1, 25, 28,
32, 34, 38n2, 53, 55, 68, 85,
85n1, 90n9, 100, 101, 105n8,
109, 120n34, 123n39, 130,
155n9, 168n40, 169, 173, 174,
175, 179, 180, 182, 212, 216,
219, 220n3, 223, 225, 238,
261, 282, 283, 287, 288, 291,
292, 309, 328, 337, 339, 340

Ewe 28, 33

F

French 2, 21, 22, 23n1, 24n2,
26, 27, 29, 31, 31n8, 32, 33,
34, 35, 36, 37, 38, 39, 41, 48,
49, 50, 86, 87, 88n6, 98n15,
99, 100, 101, 102, 105n8,
108, 108n11, 109n14, 115, 116,
117, 120n34, 122n37, 122n38,
123, 129, 130, 131, 139, 142,
144, 147n14, 152, 153, 153n1,

153n2, 154, 154n3, 155,
155n7, 156, 156n10, 156n11,
156n12, 157, 157n15, 160, 161,
162, 163, 164, 165, 165n33,
168n42, 170, 170n50, 171,
173, 174, 175, 175n58, 176n61,
177, 179, 180, 181, 182, 183,
228, 232, 233, 235, 241, 243,
246n2, 279, 283, 284, 287,
288, 293, 295, 305, 309, 312,
314n5, 316, 316n8, 316n9,
317, 317n12, 320, 321, 322,
323, 339

G

German 17, 18, 30, 33, 105n8,
109n13, 112, 123n40, 178,
181, 217, 220n3, 223, 263n2,
291, 339

Germanic 33, 36, 55, 67, 109,
123n39, 157, 157n16, 178, 279

Greek 26, 105n8, 108n11, 110,
130, 263n1, 263n2, 275n26,
276, 277, 278, 315, 315n7

H

Hebrew 18, 108n11, 263n2,
275n26, 317n10, 317n13

Hungarian 155n9, 183, 218, 219

I

Ibero-Romance 307

Icelandic 21, 22, 23, 28, 32, 33,
34, 117n25

Irish 105n8, 313, 317, 318, 319,
321, 323

Italian 3, 33, 54, 55, 61, 62,
66, 67, 68, 69, 70, 71, 78,
79, 83, 84, 104, 105n8, 108,
109n14, 117, 118n28, 118n29,
119, 120, 121, 121n36, 122,
123, 124, 124n43, 125, 126,
127, 127n44, 128, 130, 132,

156n10, 156n11, 159n22, 160,
161n27, 162, 172n54, 174, 175,
175n59, 176n61, 177, 178, 179,
180, 181, 182, 186, 188, 189,
190, 191, 198, 205, 209, 211,
238, 239n4, 242, 243, 245,
247, 251, 252, 253, 254, 255,
256, 261, 262, 263n2, 264,
269, 276, 277, 291, 292, 293,
295, 305, 322

Italo-Romance 3, 53n1, 103,
105, 106, 107, 108n11, 122,
127, 128, 239, 239n4, 244
see also Deliano, Italian,
Marsalese, Neapolitan,
Padua, Pagotto,
Piedmontese, Sicilian
dialects, Viozene

J

Japanese 29, 32, 144, 144n13,
152, 181

L

Latin 3, 22, 58, 58n4, 59, 104,
109, 114, 117, 130, 236, 239n4,
241, 243, 260, 263, 263n2,
264, 265, 266, 266n8, 267,
267n12, 270, 271n22, 271n23,
272, 275n26, 276, 277, 278
Lebanese Arabic 314, 317n10,
317n13, 321, 322

M

Mandarin 21, 22, 28, 30, 32
Marsalese 53, 54, 55, 56, 57,
64, 66

N

Neapolitan 232, 233, 234,
234n1, 235, 236, 236n3, 237,
238, 241, 247n3, 254

P

Paduan 119, 120, 156n10,
172n52
Pagotto 84, 86n2, 101
Piedmontese 172n52
Portuguese 1, 3, 100, 105n8,
108n11, 109n14, 116, 122n38,
135, 136, 136n2, 137, 137n3,
138, 140, 140n7, 142, 143,
144, 148, 150, 151, 152,
161n26, 185, 199, 293, 295,
305, 325, 326, 328, 331, 334,
338, 339, 340

E

early Romance 240

M

Medieval Romance 239

R

Romanian 3, 105n8, 109n14,
115, 115n20, 116, 117, 124n42,
130, 215, 216, 217, 218, 219,
220, 224, 227, 228, 229, 233,
293, 295, 303, 305

Russian 11, 18, 105n8, 170n50

S

Selayarese 313, 317, 318, 319,
320, 321, 323

Sicilian dialects 3, 53, 67

Slavic languages 90, 90n9,
228

Spanish 1, 2, 3, 5, 6, 7, 9, 17, 18,
50, 85, 86, 86n2, 87, 88, 89,
90, 90n9, 91, 91n11, 92, 93,
94n12, 98, 99, 100, 101, 102,
105n8, 108n11, 109n14, 116,

117, 119n32, 120n33, 120n34,
121, 122, 122n38, 123, 127,
128, 147n15, 152, 155, 155n8,
159n22, 164n31, 165, 165n32,
171, 172n54, 177n62, 178, 181,
183, 185, 186, 189, 190, 191,
191n7, 193, 198, 199, 212, 245,
248, 249, 249n7, 251, 252,
255, 256, 257, 258, 259, 260,
261, 262, 286, 293, 295, 299,
300, 302, 306, 307, 325, 326,
328, 334, 338, 339, 340

Swedish 49, 53, 55

V

Viozene 233, 234, 235

W

Welsh 105n8, 317n13

Subject index

A

agreement 2, 61, 62, 67,
114n19, 118, 133, 151, 153,
154, 154n3, 155, 157, 157n16,
157n17, 159, 159n22, 162,
162n29, 163, 164, 164n31,
165, 165n33, 166, 166n36,
170n49, 173, 174n57, 175,
176, 177, 178, 179, 180, 181,
182, 183, 187, 188, 189, 190,
197, 198, 243, 278, 280, 281,
288, 290, 317n13, 318, 319,
324, 337
argument structure 54,
249n7, 261, 263, 267, 277,
278, 282, 283

B

binding
anaphor 2, 21, 22, 27, 28,
28n6, 32, 33
anaphora 33, 34, 201, 202n1,
204, 205, 213, 323
long distance binding 2, 33
bridging 1, 201, 202, 202n1,
203, 204, 205, 206, 207, 208,
209, 210, 211, 212

C

causativization 2, 245, 246n2,
247, 251, 252, 258
clitics 2, 121, 121n35, 154, 155,
155n7, 156n12, 161, 163, 164,
165, 170n48, 178, 179, 181,
182, 183, 231, 232, 233, 234,
235, 236, 236n3, 237, 238,
239, 241, 242, 243, 244, 287,
316n8, 323
locative clitic 245, 254, 255,
256, 257, 260
Condition A 21, 22, 23, 28,
28n6, 29, 30, 31, 32

control 2, 40, 41, 42, 133, 151,
179, 183, 199, 209, 279, 280,
281, 282, 283, 284, 285, 286,
288, 289, 290, 291, 292
counter-expectational 3, 69,
70, 71, 72, 76, 77, 78, 81, 82

D

demonstrative 153, 168,
168n40, 169, 170, 171, 173,
175, 176, 177, 177n64
denominals 263, 264, 265,
266, 266n7, 276
dependent indefinites 3, 215,
216, 224, 228, 229
deponents 3, 263, 264, 264n3,
265, 266, 266n7, 267n13,
271, 271n23, 272, 276, 277
discourse 3, 21, 26, 27, 28, 29,
33, 34, 38, 69, 71, 74, 78, 80,
81, 81n16, 82, 83, 84, 100,
111, 127n45, 152, 187, 192, 198,
199, 201, 202, 203, 204, 205,
210, 211, 212, 213, 306
dislocation 1, 2, 71, 78, 84,
142, 143n11, 149, 156, 159n22,
167, 167n38, 170, 171, 172,
181, 201, 202, 203, 205, 206,
207, 208, 209, 210, 211, 212,
309, 313n4, 314n6, 317, 323
clitic left dislocation 71, 78,
84, 159n22, 309, 313n4
left dislocation 2, 71, 78, 84,
142, 159n22, 167, 201, 309,
313n4, 323
right dislocation 1, 142,
143n11, 149, 170, 201,
203, 211
distributive operator 215, 216,
218, 219, 222, 226, 227, 228
doublets 1, 135, 136, 150
doubling 2, 50, 62, 101,
143n11, 153, 154, 155, 155n9,

163, 164, 164n31, 165, 165n33,
166, 167, 170, 171, 172n52,
172n53, 172n54, 176, 177, 178,
180, 181, 182, 183, 235, 237,
244, 261, 311, 312, 319
duration 3, 85, 93, 97, 98, 325,
326, 327, 328, 329, 330, 331,
332, 333, 333n6, 334, 335, 336,
337, 337n7, 338, 339

E

echo questions 85, 86n3, 87,
88, 89, 90, 91, 92, 97, 101,
150

F

focus 7, 26, 47, 49, 50, 78, 79,
80, 82, 83, 86n3, 89, 92, 97,
99, 102, 140, 149, 151, 152,
188, 199, 206, 209, 211, 212,
239, 246, 265, 299, 303, 309,
317n13, 318, 319, 323, 325,
326, 328, 329, 330, 330n4,
331, 332, 333, 333n6, 334, 335,
336, 337, 338

G

generic *pro* 185, 186, 190, 193,
194n9, 195, 195n10, 196,
197, 198
generic *uno* 2, 185, 186, 190,
191, 192n8, 193, 194, 194n9,
195, 198
genericity 21, 24, 25n4, 27,
280, 282
genitives 103, 104, 114, 114n19,
115, 116n24, 123n40, 130

I

impersonal *se* 185, 186, 193,
198, 277
implicit arguments 128, 279,
280, 282, 289, 291, 292

- inflected construction 3, 53,
 53n1, 55, 56, 67
 information-seeking questions
 1, 85, 94, 95, 97, 98, 98n15
 intonation 1, 50, 70, 71, 71n3,
 81, 82, 85, 86n3, 87, 90, 92,
 94, 95, 98, 100, 138, 138n6,
 139, 144, 170, 207n3, 297,
 299, 299n2, 300, 301, 339,
 340
 inversion 2, 136, 142, 150, 152,
 153, 153n2, 154, 155, 161, 177,
 179, 180, 181, 182, 183
 island 2, 139, 309, 312, 313, 314,
 314n6, 315, 316, 317, 319, 320,
 321, 322, 323
 isochrony 325, 326, 336, 337,
 338
- K**
- kinship nouns 103, 104,
 118n28, 121n35, 128
- L**
- left periphery 69, 78, 79, 83,
 84, 101, 142, 152, 182, 187,
 287, 292, 298, 305, 306, 307,
 315, 321, 324
 logophoricity 2, 21, 22, 28,
 31n8, 33, 34, 279, 280, 282,
 285, 291
- M**
- Middle 263, 263n2, 264,
 264n3, 265, 266, 268n16,
 269, 270, 271, 275, 275n26,
 276
 motion verbs 53, 54, 56, 57,
 64, 67, 245
 movement 2, 85, 88, 88n6, 89,
 99, 100, 101, 102, 108, 109,
 110, 114n19, 125, 129, 131, 135,
 138, 144, 144n13, 145, 150,
 151, 153, 157, 158, 158n20,
 159n22, 160, 161, 161n27,
 162, 162n30, 163, 168n42,
 170n50, 173n55, 177, 179,
 180, 182, 238, 242, 248n5,
 256, 278, 280, 291, 299, 307,
 309, 310, 311, 312, 312n2, 313,
 314, 314n5, 314n6, 315, 316,
 317, 318, 319, 320, 321, 322,
 322n14, 323, 324
- N**
- negation 2, 5, 6, 11, 12, 13, 16,
 18, 19, 35, 36, 38, 38n2, 39,
 40, 49, 50, 51, 69, 70, 82, 83,
 131, 304
 double negation 2, 35, 36,
 39, 50
 expletive negation 2, 5,
 18, 69
 negative concord 2, 18, 35,
 35n1, 36, 49, 50, 51
 noun phrase 103, 129, 130, 183
 null subjects 179, 185, 186,
 189, 196, 197, 198, 199
- O**
- optatives 1, 293, 295, 296, 298,
 299, 300, 301, 302, 303, 304
- P**
- Parametric Comparison
 Method 103, 131
 partition 36, 68, 128, 215, 218,
 220, 222, 224, 227
 passive 2, 179, 199, 263n2,
 264, 264n3, 265, 270, 276,
 279, 280, 283, 284, 285, 286,
 288, 289, 290, 291, 292
 person 2, 21, 22, 24, 25, 26,
 27, 28, 31n8, 32, 33, 40,
 54, 62, 107, 108, 126, 127,
 127n44, 128, 132, 133, 153,
 159n22, 164, 164n31, 165,
 165n32, 166, 166n36, 167,
 169, 169n47, 170, 170n49,
 170n50, 171, 172n52, 172n54,
 173, 174, 175, 175n58, 176,
 176n61, 177, 177n64, 180,
 185, 187, 188, 188n4, 188n5,
 190, 193, 197, 199, 201, 207,
 236, 241, 243, 280, 283,
 299, 307
 possessives 103, 104, 116, 117,
 117n27, 118, 118n28, 118n31,
 119, 120, 120n33, 121, 121n35,
 121n36, 122, 122n38, 123,
 123n39, 123n40, 124, 124n42,
 124n43, 125, 126, 131
 prosody 2, 35, 36, 38, 39, 48,
 49, 50, 51, 71, 100, 102, 231,
 299, 337
- R**
- reconstruction 48, 131, 309,
 310, 311n1, 312, 314, 316, 321,
 322, 323, 324
 restructuring verbs 53, 54,
 55, 56
 resumption 2, 235, 309, 310,
 311, 312, 313, 314n5, 316, 317,
 317n13, 319, 322, 323, 324
 rhythm 325, 326, 327, 328,
 333n6, 335, 336, 337, 338,
 339, 340
- S**
- semantics 1, 2, 3, 5, 6, 7, 10, 11,
 16, 18, 19, 33, 34, 49, 50, 61,
 83, 84, 89, 101, 183, 199, 211,
 212, 213, 216, 221, 228, 229,
 266n8, 269, 273, 278, 292,
 297, 300, 306, 323, 324
 stress 2, 85, 89, 90, 92, 93, 95,
 97, 98, 100, 138, 144, 209,
 231, 232, 233, 234, 234n1,
 235, 236, 236n3, 241, 242,
 243, 244, 325, 326, 327, 328,
 329, 330, 331, 332, 333, 334,
 335, 336, 337, 338, 339, 340
 marked stress 85, 95
 stress-timing 325, 326, 327,
 328, 338, 339
 subjunctive 23, 27, 57, 72, 175,
 293, 294, 295, 297, 297n1,
 300, 303, 303n4, 304, 305,
 306, 307
 surprise questions 1, 69, 71,
 84, 85, 90n9, 91, 92, 94, 95,
 97, 98
 syllable-timing 325, 327, 337,
 339, 340
 syntactic microvariation 245
 syntax 1, 2, 19, 33, 34, 37, 50,
 53, 67, 68, 69, 71, 71n3, 83,
 84, 87, 88, 89, 99, 100, 101,
 102, 104, 105, 120n34, 130,
 131, 132, 151, 152, 157n15, 178,
 179, 180, 181, 182, 183, 187,
 199, 211, 212, 229, 231, 232,
 235, 243, 244, 252n9, 261,
 262, 263, 273, 276, 277, 278,
 280, 287, 288, 291, 292, 306,
 310, 318, 323, 324

T

temporal adverbial 5, 17
tonal range 85, 92, 95, 96, 98
topic-identification 185, 187,
190, 196, 198

V

Visser's Generalization 279,
284, 292

W

wh-in-situ 1, 85, 85n1, 86, 87,
88, 89, 90, 92, 93, 97, 98,
100, 101, 102, 136, 136n2,
138, 139, 140, 141, 144, 145,
146, 150, 151
wh-movement 2, 99, 100, 101,
135, 138, 144, 144n13, 145,
150, 151, 309, 314n6, 321

wh-questions 1, 84, 92, 95,
97, 100, 101, 135, 135n1, 137,
137n4, 140, 141, 142, 147,
148, 150, 151, 313, 313n4, 314,
318, 321

This book contains a peer-reviewed selection of papers presented at the 46th Linguistic Symposium on Romance Languages (LSRL 46) that took place in April 2016 at Stony Brook University (SUNY), New York. The most current research and debates on bilingualism, historical linguistics, morphology, phonology, semantics, sociolinguistics, and syntax can be found in its pages. This collection will be of interest to Romance linguists and general linguists as well.



John Benjamins Publishing Company