

Current Issues in Linguistic Theory

# Contemporary Approaches to Romance Linguistics

EDITED BY

Julie Auger

J. Clancy Clements

Barbara Vance

# CONTEMPORARY APPROACHES TO ROMANCE LINGUISTICS

AMSTERDAM STUDIES IN THE THEORY AND  
HISTORY OF LINGUISTIC SCIENCE

General Editor

E.F. KONRAD KOERNER

(Zentrum für Allgemeine Sprachwissenschaft, Typologie  
und Universalienforschung, Berlin)

Series IV – CURRENT ISSUES IN LINGUISTIC THEORY

*Advisory Editorial Board*

Lyle Campbell (Salt Lake City); Sheila Embleton (Toronto)  
Brian D. Joseph (Columbus, Ohio); John E. Joseph (Edinburgh)  
Manfred Krifka (Berlin); E. Wyn Roberts (Vancouver, B.C.)  
Joseph C. Salmons (Madison, Wis.); Hans-Jürgen Sasse (Köln)

Volume 258

Julie Auger, J. Clancy Clements and Barbara Vance (eds)

*Contemporary Approaches to Romance Linguistics.*  
*Selected Papers from the 33rd Linguistic Symposium on Romance Languages*  
*(LSRL), Bloomington, Indiana, April 2003.*

CONTEMPORARY  
APPROACHES TO ROMANCE  
LINGUISTICS

SELECTED PAPERS FROM  
THE 33RD LINGUISTIC SYMPOSIUM  
ON ROMANCE LANGUAGES (LSRL),  
BLOOMINGTON, INDIANA, APRIL 2003

Edited by

JULIE AUGER  
J. CLANCY CLEMENTS  
BARBARA VANCE

*Indiana University, Bloomington*

JOHN BENJAMINS PUBLISHING COMPANY  
AMSTERDAM/PHILADELPHIA



TM

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

#### Library of Congress Cataloging-in-Publication Data

Linguistic Symposium on Romance Languages (33rd : 2003 : Bloomington, Ind.).

Contemporary approaches to Romance linguistics : selected papers from the 33rd Linguistic Symposium on Romance Languages (LSRL), Bloomington, Indiana, April 2003 / edited by Julie Auger, J. Clancy Clements, Barbara Vance.

p. cm. -- (Amsterdam studies in the theory and history of linguistic science. Series IV, Current issues in linguistic theory, ISSN 0304-0763 ; v. 258)

Includes bibliographical references and index.

Romance languages--Congresses.

PC11 .L53 2003

440--dc22

2004052953

ISBN 90 272 4772 2 (Eur.) / 1 58811 598 4 (US) (Hb; alk. paper)

© 2004 – 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 Co. • P.O.Box 36224 • 1020 ME Amsterdam • The Netherlands  
John Benjamins North America • P.O.Box 27519 • Philadelphia PA 19118-0519 • USA

## CONTENTS

Case, Agreement, and Expletives: A parametric difference in Old French and Modern French <i>Deborah Arteaga &amp; Julia Herschensohn</i>	1
Paradigmatic and Syntagmatic Relations in Italian Verbal Inflection <i>Luigi Burzio</i>	17
The Role of the L1 in the Overgeneralization of Causatives in L2 English and L2 Spanish <i>Mónica Cabrera &amp; María Luisa Zubizarreta</i>	45
The Inchoative Interpretation of the <i>Imperfecto</i> <i>Alicia Cipria</i>	65
Reinterpreting the CV Transition: Emergence of the glide as an allophone of the palatal lateral <i>Laura Colantoni</i>	83
Intervocalic Velar Nasals in Galician <i>Sonia Colina</i>	103
Null Objects in French and English <i>Sarah Cummins &amp; Yves Roberge</i>	121
Micro-Parametric Variation and Negative Concord <i>Viviane Déprez &amp; France Martineau</i>	139
Contrast and Addition in Romance: A case study in microvariation <i>Luis Eguren &amp; Cristina Sánchez</i>	159
On the Structure of Syncretism in Romanian Conjugation <i>Ronald F. Feldstein</i>	177
Sluicing in Romanian: A typological study <i>Frederick Hoyt &amp; Alexandra Teodorescu</i>	197

Romance Intonation From a Comparative and Diachronic Perspective: Possibilities and limitations <i>Jose Ignacio Hualde</i>	217
Romance SE as an Aspectual Element <i>Paula Kempchinsky</i>	239
Proto-Romance *[w] and the Velar Preterites <i>Eric Lief</i>	257
The Phonological Role of Paradigms: The case of insular Catalan <i>Maria-Rosa Lloret</i>	275
A Constraint-Based Analysis of Galician <i>Geda</i> <i>Fernando Martínez-Gil</i>	299
Peak Placement in Two Regional Varieties of Peruvian Spanish Intonation <i>Erin O'Rourke</i>	321
On Intentional Causation in Italian <i>Antonella Vecchiato</i>	343
Inversion, Reconstruction, and the Structure of Relative Clauses <i>Luis Vicente</i>	361
Infinitival Complement Constructions in Spanish: A Construction Grammar approach <i>Jiyoung Yoon</i>	381
Index of Terms and Concepts	399

## PREFACE

The articles that follow represent a selection of the papers presented at the 33rd annual meeting of the Linguistic Symposium on Romance Languages held at Indiana University, Bloomington, on April 24-27, 2003. Each article submitted for consideration was reviewed by three anonymous reviewers, whom we would like to thank for their careful reading and helpful comments. We would also like to express our deepest appreciation to our assistant editor, Rachel Thyre Anderson, who lent her expertise in both Romance linguistics and copyediting to the preparation of the final camera-ready manuscript. For financial support, we are grateful to the following units at Indiana University: the Office of the Vice President and Chancellor, the Vice President for Research and Dean of the University Graduate School, the Department of French and Italian Mary-Margaret Barr-Koon Fund, the Department of Spanish and Portuguese, the College of Arts and Sciences, the Office of International Programs, and the Department of Linguistics. Finally, neither the conference nor this volume would have been possible without the tireless work of Deborah Piston-Hatlen, who provided the organizational structure for the entire project stretching over nearly three years.

A name that belongs to this volume in spirit, if not in terms of direct responsibility for its content, is that of Albert Valdman. During his 44 years at Indiana University, Professor Valdman has inspired hundreds of students and colleagues with his careful scholarship, his interactive teaching style, his innovative guidance on pedagogical matters, his ability to organize large teams of researchers into efficient intellectual units, and his constructive and cooperative approach to every issue—large or small—that arises concerning the study and teaching of linguistics and French at Indiana University. Professor Valdman was one of the first scholars to contribute to LSRL, publishing papers in the first, fifth, and seventh volumes of proceedings as well as in the proceedings of LSRL 3, held here in Bloomington and co-organized (with Mark Goldin) by our colleagues Joe Campbell and Mary Clayton. After helping to organize LSRL 33 and two other more specialized conferences that met in Bloomington immediately before LSRL, Albert Valdman officially retired from Indiana University in May 2004. With gratitude for his countless contributions to Romance Linguistics and his continuing leadership here and around the world, we dedicate this volume to him.

Bloomington, July 2004

Julie Auger  
J. Clancy Clements  
Barbara Vance





**CASE, AGREEMENT, AND EXPLETIVES**  
**A PARAMETRIC DIFFERENCE IN OLD FRENCH AND MODERN FRENCH\***

DEBORAH ARTEAGA & JULIA HERSCHENSOHN  
*University of Nevada, Las Vegas & University of Washington*

**0. Introduction**

The investigation of expletives cross-linguistically has been pivotal in recent research (Chomsky 2000, 2001a, 2001b) because they are, in a sense, manifestations of pure syntax, virtually devoid of meaning yet satisfying requirements of EPP and Case. In this paper we re-examine expletives in Old (OF) and Modern French (MF), languages that differ parametrically with respect to the distribution of expletives and Case agreement. We begin our discussion by examining the contrasting distribution of expletives in the two languages, focusing largely on subject-verb agreement patterns and Case assignment. After reviewing the previous analysis of Arteaga (1994), we present the relevant theoretical assumptions under the Minimalist Program. Finally, we give our analysis of expletive constructions in OF and MF. We propose that neuter *il* merges at CP only to satisfy the EPP of C.

**1. Expletives in OF and MF**

Like many other Romance languages, OF allowed null subjects, as illustrated in (1), where the first-person subject is unexpressed:<sup>1</sup>

---

\* We wish to thank the audiences of the University of Washington colloquium and LSRL, as well as our three anonymous reviewers, for comments and suggestions, especially Barbara Vance, Steve Franks, and Viviane Déprez. We thank Barbara for extended discussion of OF data and theory. This article is part of a larger project on expletives (cf. Arteaga & Herschensohn 2001, 2003).

<sup>1</sup> The abbreviations we use in this work are as follows: 1SG/PL (first-person singular/plural); 2SG/PL (second-person singular/plural); 3SG/PL (third-person singular/plural); EX (expletive); FUT (future); PRES (present); and INDIC (indicative). In addition, as OF had a two-case declension system, we indicate the case on nouns by the abbreviations NOM (nominative case) and OBL (oblique case), with the designations SG for singular and PL for plural. We indicate gender by M for masculine, F for feminine, and N for neuter. In MF examples, we use the abbreviation PART to indicate partitive case. Where we have culled OF examples cited by other authors, the translation is Arteaga's, unless otherwise noted.

- (1) *Or revendrai al pedre ed a la medre*  
 now come-back<sub>FUT.1SG</sub> to-the<sub>M.SG.OBL</sub> father<sub>M.SG.OBL</sub> and to the<sub>F.SG.OBL</sub> mother<sub>F.SG.OBL</sub>  
 “Now I will come back to the father and to the mother.”  
 (*La vie de Saint Alexis* 101)

Nevertheless, as noted by Adams (1987), OF is not as rigorous a null-subject language as Spanish or Italian, because, among other reasons, null subjects are quite rare in embedded clauses in OF; moreover, Roberts (1993) shows that, unlike Spanish and Italian, which are argued not to project Spec IP (Alexiadou & Anagnostopoulou 2001), OF did project that subject position. As the example in (1) illustrates, OF verbal inflection was rich, showing distinct person and number, and also showing overt VP-internal subjects, although the degree of richness of inflection is open to some debate (cf. Foulet [1919] 1995; Roberts), suggesting that OF was ripe for the parametric change that followed. These differences notwithstanding, for the purposes of our paper, we continue to refer to OF as a null-subject language.

Case marking on OF nominal elements (nouns, pronouns, adjectives, determiners) was morphologically overt, indicating gender, number, and Case (nominative or oblique): the examples in (2) and (3) provide the regular declension of singular adjectives using the adjective *bon* “good” and third-person pronouns in OF:

- (2) “good”

	Masculine	Neuter	Feminine
NOM.SG	<i>bons</i>	<i>bon</i>	<i>bone</i>
OBL.SG	<i>bon</i>	<i>bon</i>	<i>bone</i>
NOM.PL	<i>bon</i>	--	<i>bones</i>
OBL.PL	<i>bons</i>	--	<i>bones</i>

- (3)

	Masculine	Neuter	Feminine
NOM.SG	<i>il</i> “he”	<i>il/el</i> “it”	<i>ele</i> “she”
OBL.SG	<i>lo/le</i> “it”	<i>lo/le</i> “it”	<i>la</i> “it”
NOM.PL	<i>il</i> “they”	--	<i>eles</i> “they”
OBL.PL	<i>eux</i> “them”	--	<i>eles</i> “them”

As indicated in (2), diagnostic marking is seen in the masculine, where the terminal *-s* in M.SG.NOM and M.PL.OBL contrasts with  $-\emptyset$  in M.SG.OBL and M.PL.NOM, (e.g., *bons* vs. *bon*). The crucial difference between masculine and neuter adjectival forms can in turn be seen in the nominative singular (*-s* for M.SG,  $-\emptyset$  for N.SG), *bons*, masculine, *bon*, neuter. Within the pronominal system, as in (3), the masculine nominative singular, *il*, does not show the

characteristic *-s* of the M.SG.NOM, so that it is identical in form to both the masculine nominative plural and neuter nominative singular.

Another salient characteristic of OF was its obligatory V2 word order (4) in main clauses (Adams 1987; Arteaga 1994; Roberts 1993; Vance 1997), as shown in (4a), where the verb *dist* must occupy the second position:

(4) V2 in main clauses:

- a. *Dont dist li dus au chevalier*  
 therefore said<sub>3SG</sub> the<sub>M.SG.NOM</sub> duke<sub>M.SG.NOM</sub> to-the<sub>M.SG.OBL</sub> knight<sub>M.SG.OBL</sub>  
 “Therefore the duke said to the knight.”  
 (*Chastelaine de Vergi* 217, Foulet [1919] 1995:§451)
- b. *Par Petit Pont sont en Paris entré*  
 by Little<sub>M.SG.OBL</sub> Bridge<sub>M.SG.OBL</sub> were<sub>3PL</sub> in Paris entered<sub>M.PL.NOM</sub>  
 “They entered Paris by the Petit Pont.”  
 (*Charroi de Nimes* 26-31, Roberts 1993:95)
- c. *Li cuents Guillelmes fu mol gentiz et ber*  
 the<sub>M.SG.NOM</sub> count<sub>M.SG.NOM</sub> William<sub>M.SG.NOM</sub> was<sub>3SG</sub> very kind<sub>M.SG.NOM</sub> and good<sub>M.SG.NOM</sub><sup>2</sup>  
 “Count William was very kind and good.”  
 (*Charroi de Nimes* 26-31, Roberts 1993:95)

As in other V2 languages, what appears in OF to be a surface S-V-O order, as in (4c), is in fact the result of the lexical subject appearing in Spec CP while the verb is in C, thereby conforming to V2 requirements, as shown convincingly by Roberts and Adams.

The syntactic status of subject pronouns in Old French is also relevant for our discussion. It is well known that subject pronouns in OF are not clitics in all positions (Adams 1987; Roberts 1993; Vance 1997), whereas they are necessarily attached to a verb in MF. Roberts (1993:153-154) provides extensive examples in support of his claim that pronouns in OF had the following characteristics not associated with clitic pronouns: they could be stressed, they could be conjoined with other pronouns, they could be separated from the verb, they could stand alone, and they could be modified. These tonic pronouns could then occupy first position to the verb in second position. Roberts (1993:177-186) argues that enclitic subject pronouns (induced by inversion, for example) are, on the other hand, clitics; he proposes that they are cliticized from Spec IP to the verb in C, an analysis also supported by Vance.

Arteaga (1994) notes that although OF is a null-subject language, it nonetheless evinces overt expletive pronouns. She argues that the expletives were necessary to maintain V2 word order. The example in (5) shows the expletive pronoun *il* in OF, which is not expressed in (6):

<sup>2</sup> As an irregular imparisyllabic, *ber* does not show the characteristic nominative singular *-s*.

- (5) *Il est écrit es cartes*  
 EX is<sub>3SG</sub> written<sub>N.SG.NOM</sub> in-the<sub>F.PL.OBL</sub> letters<sub>F.PL.OBL</sub>  
 “It is written in the letters.”  
 (*La chanson de Roland* 1039, Jensen 1990:133)
- (6) *Et certes mout m’ est bel que vos estes li  
 plus cortois de nos*  
 and certainly very-much to-me is<sub>3SG</sub> pleasing<sub>N.SG.NOM</sub> that you<sub>2PL</sub> are<sub>2PL</sub> the<sub>M.SG.NOM</sub>  
 most courtly<sub>M.SG.NOM</sub> of us<sub>1PL.OBL</sub>  
 “Certainly, it pleases me that you are the most courtly one of us.”  
 (*Yvain* 73-74, De Lage 1975:37)

Although the subject is unexpressed in (6), it is nonetheless clear that the expression is an impersonal one from the morphological marking of the predicate nominative adjective *bel*, which, lacking the flexional *-s* of the nominative masculine singular, is neuter singular (cf. (2)).

From the examples in (1) through (6), we see that OF, a null-subject language with V2 effects in main clauses, had overt expletive pronouns. MF, on the other hand, requires overt expression of all subject pronouns, including expletives, as illustrated in (7) through (9):

- (7) *Or je reviendrai au père et à la mère*  
 Now I come-back<sub>FUT.1SG</sub> to-the father and to the mother  
 “Now I will come back to the father and to the mother.” (= 1)
- (8) *Il est écrit dans les lettres*  
 EX is<sub>3SG</sub> written in the letters  
 “It is written in the letters.” (= 4)
- (9) *Certes, il m’ est bien agréable que vous soyez si cortois*  
 certainly it to-me is<sub>3SG</sub> very pleasant that you<sub>2PL</sub> are<sub>2PL</sub> so courtly  
 “Certainly, it is pleasant to me that you are so courtly.” (= 5)

MF lacks V2 effects, meaning that it only allows XP fronting to CP in interrogatives or topicalized sentences. Compare (10), where the verb and subject are in TP, with (4):

- (10) *Donc le duc dit au chevalier*  
 therefore the duke said<sub>3SG</sub> to-the knight  
 “Therefore the duke said to the knight” (= 4)

Another striking difference between OF and MF is subject-verb agreement in matrix expletive constructions. As noted by Arteaga (1994), OF shows non-uniformity of agreement in expletive constructions, as illustrated in (11) and (12):

- (11) *Il ot plusors qui burent a outrage.*  
 EXhad<sub>3SG</sub> several<sub>M.PL.OBL</sub> who drank<sub>3PL</sub> to excess<sub>M.SG.OBL</sub>.  
 “There were several who drank to excess.”  
 (*Roman de Troie* 95, Einhorn 1974:§188)
- (12) *Il i maneit un tyrant*  
 EXthere lived<sub>3SG</sub> a<sub>M.SG.OBL</sub> tyrant<sub>M.SG.OBL</sub>.  
 “There lived a tyrant.”  
 (*Ambroise, Guerre Sainte* 1385, Jensen 1990:4)

In (11), it appears that agreement obtains between the expletive *il* and the verb *ot*; *plusors* has oblique (partitive) Case. In (12), the postverbal associate *un tyrant* is in the oblique Case; the verb is singular, apparently agreeing in number with *il*. In (13), however, the plural verb *vont* agrees with the postverbal associate *ci vieil prestre*, which is in the nominative Case; similarly, in (14), the singular postverbal subject *ses obers blancs* “his white hauberck” is in the nominative Case; the verb is singular.

- (13) *Il i vont ci vieil prestre*  
 EXthere go<sub>3PL</sub> these<sub>M.PL.NOM</sub> old<sub>M.PL.NOM</sub> priests<sub>M.PL.NOM</sub>.  
 “There go these old priests.”  
 (*Aucassin* §111, 6, Togeby 1974:§111, 6)
- (14) *Il nel gari ses obers blancs*  
 EXnot-him protected<sub>3SG</sub> his<sub>M.SG.NOM</sub> hauberck<sub>M.SG.NOM</sub> white<sub>M.SG.NOM</sub>.  
 “His white hauberck didn’t protect him.”  
 (*Le Brut de Munich* 1775, Horning 1880:246)

Although on the surface it could appear that examples like (13) and (14) are cases of right dislocation, here we follow Arteaga (1994), who follows Horning (1880), Jensen (1990), and Moignet (1976) in arguing that the pronoun in question is the neuter *il* and not the masculine plural (*il* in OF.) In support of her claim, she notes that in cases where there is a diagnostic feminine, the pronoun expressed is invariably *il* not *ele*, as in (15a), which shows verbal agreement with the postverbal lexical subject, or (15b), which shows agreement with the postverbal enclitic *je*.

- (15) *Il/Ce* expletives with postverbal nominative:  
 a. *Li chastiaus dont il parloient tantes gens*  
 the chateau<sub>M.SG.NOM</sub> of-which EX spoke<sub>3PL</sub> many<sub>F.PL.NOM</sub> people<sub>F.PL.NOM</sub>.  
 “The castle so many people were talking about.”  
 (*Montreuil* 9312, Jensen 1990:§292)

- b. *Ce sui je*  
 this<sub>N,SG</sub> am<sub>1SG</sub> I<sub>1SG,NOM</sub>  
 “It is I.”  
 (*Queste* 250.27, Tobegy 1974:§106).

Arteaga further notes that agreement between a verb and a postverbal NP is found in other constructions in OF. Finally, in our view, the fact that neuter *il* is interchangeable in these constructions with *ce/ço*, as illustrated in (16), provides further evidence for our analysis (cf. Piatt 1898):<sup>3</sup>

- (16) *Ço peiset moi que ma fin tant demoret.*  
 It/That<sub>N,SG</sub> grieves<sub>3SG</sub> me<sub>OBL,SG</sub> that my<sub>F,SG,NOM</sub> end<sub>F,SG,NOM</sub> so-long takes  
 “It chagrins me that my end is taking so long.”  
 (*La vie de Saint Alexis*, 92e, Piatt 1898:31)

Summarizing, then, sentences (13) and (14) constitute a dilemma, since they appear to have two nominative Cases, the postverbal lexical DP and the expletive pronoun; moreover, subject-verb agreement appears to be non-uniform.

In contrast to the variable agreement found in OF expletive constructions, MF tensed clauses always show agreement with the expletive, as in (17). Here the matrix verb *risquer* agrees with the expletive, although the logical subject of unaccusative *arriver* “to arrive” under raising *risquer* “to risk, to be apt” is a long distance plural associate:

- (17) MF expletives  
 a. *Il risque (\*risquent) d’ arriver trois hommes.*  
 EX risks<sub>3SG</sub> (\*risk<sub>3PL</sub>) of to-arrive three men  
 “There are apt to arrive three men.”  
 b. *Trois hommes risquent d’ arriver.*  
 three men risk<sub>3PL</sub> of to-arrive  
 “Three men are apt to arrive.”

In MF, agreement obtains with the associate only if the associate raises to the matrix subject, as in (17b); there is no long-distance agreement in MF of the type found in Modern English (e.g., (17a)).

<sup>3</sup> An anonymous reader questions the status of *ço* as expletive in (16), suggesting that it is instead a demonstrative. In any case, *ço* still fulfills the function of occupying first position, thus supporting our claim. The issue at hand seems to be whether or not *ço* has yet been bleached of demonstrative force in a process of grammaticalization.

## 2. *Theory*

### 2.1 *Previous analyses of expletive constructions*

Arteaga (1994) argues that the nonuniformity of agreement and Case marking in OF expletive constructions can be accounted for by Roberts' (1993) proposal for variable nominative Case assignment in OF. Roberts, using GB theory, proposed that two options for nominative Case assignment existed in OF, namely, assignment under agreement or under government. In the case of OF expletive constructions, Arteaga argued that where agreement obtains between the verb and the postverbal DP, such as "the old priests" in (13), I assigns nominative Case to the VP-internal subject under government. Where the verb is third singular, as in (11), she proposed that it agrees with the expletive *il* and that I assigns nominative Case to the subject *il* in IP. Arteaga argues that the existence of the overt expletive in OF is a function of its V2 word order; following Roberts, she assumes that C contains AGR in V2 languages, triggering raising of the verb to C and *il* to Spec CP.

While the analysis in Arteaga (1994) relates, correctly, in our view, V2 word order to overt expletives, the expression of *il* in examples like (13) and (14) remains a stipulation that is not theoretically motivated. Another weakness of this approach is that the variable agreement and Case assignment in OF expletive expressions remains unmotivated, and indeed, arbitrary. In other words, why does verb agreement take place in CP or IP or VP? In the next section we review the theoretical assumptions within the Minimalist Program, which, as we argue, allow us to provide a more principled account of expletives in OF.

### 2.2 *Theoretical assumptions*

In the bare phrase structure framework of the Minimalist Program (Chomsky 2000, 2001a, 2001b) syntactic combinations result mainly from the operations Merge and Agree. Merge is in part determined by uninterpretable morphological features of functional categories, and syntactic structure is built from an array of lexical items that combine from the bottom up (Epstein & Hornstein 1999; Ndayiragije 1999). Agree is an operation that matches features and deletes uninterpretable ones; it may occur with or without overt movement. This approach eliminates the analysis proposed in Chomsky (1995) calling for covert movement to check agreement. Agreement may be checked with the elements in situ. English and MF require overt subjects and verb agreement; these syntactic phenomena are motivated respectively by the uninterpretable features [+EPP, +Agr], requiring overt interpretable features to check and delete.



Certain scholars have considered that a more economical approach to expletive checking might reduce EPP effects to Case theory. In aiming to create as economical a system as possible, Martin (1999:2) argues that within a Minimalist approach, effects of EPP features “mostly follow from independent principles” and “a significant portion of the EPP is reducible to Case.” In the same volume, Groat (1999) proposes that the expletive *there* in English checks an uninterpretable Case feature, obviating the need for the exceptional checking of EPP proposed in Chomsky (1995). However, English *there* is morphologically distinct from expletives in languages like French (cf. Arteaga & Herschensohn 2001), and can therefore not be taken as a prototype of expletive behavior.

Alexiadou and Anagnostopoulou (1998, 2001) and Holmberg (2000, 2002) have provided ample evidence that parametric variations in a range of languages other than English require the availability of uninterpretable Case, EPP, and Agr. Our present work assumes and argues further for this approach. Alexiadou and Anagnostopoulou (1998) propose that languages can satisfy the EPP of tense/agreement either by movement of XP (the sentential subject) or by a nominal category of X<sup>0</sup>, namely, the inflected verb richly endowed with explicit person and number agreement. The choice is a parametric option generally distinguishing overt- and null-subject languages respectively. Alexiadou and Anagnostopoulou (1998) note that MF is anomalous in retaining verb raising (a characteristic of null-subject languages), but requiring XP movement to satisfy EPP in T.

Holmberg (2000), in a discussion of stylistic fronting in Icelandic, extends this idea, arguing that an aspect of the EPP feature of I, which he calls the [P] feature, requires that the Spec of IP be lexically filled in sentences with a finite verb. Holmberg (2000:456) says the EPP embodies two separate features of I, which must be satisfied, D and P. He explains this “double” checking by stating that “[D] in I attracts all available instantiations of D in its domain.” He argues that the P feature is a phonological feature [p-Feature], which he adds to formal syntactic and semantic features. He claims that syntactic categories can “see and operate on” P features. Crucially, it is the p-Feature matrix of the category that moves, and must move to the Spec of IP, regardless of its status as head or XP. This accounts, in his view, for the fact that stylistic fronting in Icelandic affects a range of categories, and for the lack of focus effects.

In such a framework, MF T is [+EPP, +Agr], so merger of the overt subject to T is necessary, as is checking of verbal agreement.<sup>4</sup> In OF, on the other hand, rich nominal marking on the verb satisfies the EPP of T, and no overt nominative subject is required. OF, as a V2 language, requires raising of an XP to Spec CP; we extend Holmberg's (2000) proposal to CP, arguing that it is the EPP feature of C that must be deleted by fronting of an XP. XP can be of any category (AP, DP), but must be phrasal. We assume both T and C in OF also have an uninterpretable [+Agr] feature that is satisfied by verb raising to T and C. We use [+Agr] to represent finer feature(s) satisfied by the raised verb. T requires finiteness, whereas C may check mood. While the details are outside the scope of our study, Roberts (1993) has argued for the feature [+Agr] in C of V2 languages. We assume that head movement of the verb is a PF phenomenon (Chomsky 2001a, 2001b; Boeckx & Stjepanovic 2001), but that the appropriate features are checked in vP, TP, and CP.<sup>5</sup>

### 3. *Analysis*

As we saw in the first section, MF expletive constructions differ from those in OF in that the verb in MF must agree with the expletive subject *il*, never with the long-distance associate. In OF, however, there is apparently variable agreement, either with the impersonal 3SG verb, as in (11) and (12), or postverbal DP, as in (13) and (14). Our analysis derives the diachronic differences from the morphological differences between the two languages by assuming that the rich inflection of OF provided interpretable features with the ability to delete a broader range of uninterpretable features than in MF.

The analysis of MF expletive constructions is relatively straightforward. Consider the MF examples in (18) and (19):

- (18) *Il arrive des jeunes filles.*  
 EXarrive<sub>3SG</sub> some young girls<sub>F,PL,PART</sub>  
 "Some young girls arrive."
- (19) *Les jeunes filles arrivent.*  
 the young girls<sub>F,PL,NOM</sub> arrive<sub>3PL</sub>  
 "The young girls arrive."

<sup>4</sup> Koeneman and Neeleman (2001:215) classify verbs according to rich inflection and V to I, as follows: Poor (No V to I, no pro drop), Middle Class (V to I, no pro drop), and Rich (V to I, pro drop). In this schema MF is Middle Class.

<sup>5</sup> While the exact formulation of verb raising is not relevant to our argument, it is possible that head movement of the verb is a PF phenomenon.

The array for (18) contains expletive *il* (marked for gender, number, and Case), the 3SG verb, an associate marked for partitive Case, and T with [+EPP, +Agr]. The associate *filles* is checked at the vP phase to eliminate uninterpretable partitive Case.<sup>6</sup> At TP, expletive *il* [3SG.NOM] merges with v+T, which has interpretable [3SG.PRESENT] features. Interpretable features are matched against one another, and they also check and delete the uninterpretable EPP, Agr, and nominative. In (19) the associate is forced to move to subject position to check off [+EPP, +Agr] as in the preceding scenario. Any other Case or tense array would crash (e.g., two nominatives) because there would be uninterpretable features left unchecked and violating Full Interpretation.

What happens in OF matrix clauses? As we have seen, two scenarios are possible with expletive *il*, the associate with nominative Case (13) and (14), or with oblique Case (11) and (12), repeated here as (20) and (21):

- (20) *Il i vont ci vieil prestre*  
 EX there go<sub>3PL</sub> these<sub>M,PL,NOM</sub> old<sub>M,PL,NOM</sub> priests<sub>M,PL,NOM</sub>  
 “There go these old priests.” (= 13)
- (21) *Il ot plusors qui burent a outrage.*  
 EX had<sub>3SG</sub> several<sub>M,PL,OBL</sub> who drank<sub>3PL</sub> to excess<sub>M,SG,OBL</sub>  
 “There were several who drank to excess.” (= 11)

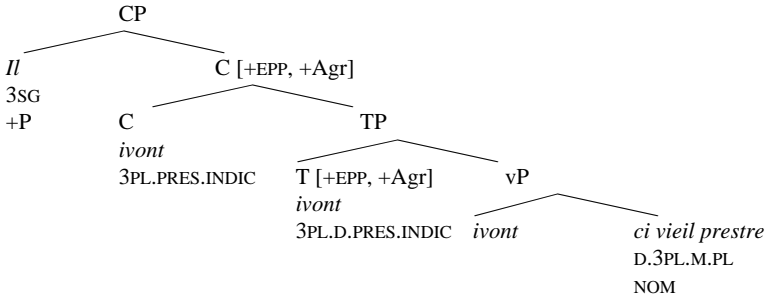
Given our theoretical assumptions, no longer are we obliged to see (20) and (21) as structurally different. In both sentences, EPP and Agr of T are satisfied by rich verbal inflection. In (20) the subject “old priests” has its nominative Case checked by matching verb agreement. Extending Arteaga’s (1994) idea of neuter *il* as CP expletive, we propose that *il* is merged only at the CP phase to satisfy the EPP of C in both (20) and (21). It is not checked for Case at this level, and nominative is irrelevant. CP expletive *il* serves the same function as Caseless *ce/ço* “that,” which can appear in any sentential position (cf. Piatt 1898).

To illustrate, in the array (22), {*il*<sub>3SG,N</sub>, *vont*<sub>3PL</sub>, *prestre*<sub>3PL,M,NOM</sub>, T [+EPP, +Agr], C [+EPP, +Agr]}, the nominative associate requires merge/agree at TP without movement. Because the verbal inflection in OF is nominally robust, no movement to Spec TP of the postverbal DP is required. Interpretable tense of *vont* deletes [+Agr] of T, and the nominal interpretable feature (e.g., D) of the rich inflection deletes [+EPP] of T. The nominative Case of *prestre* is deleted through Agree with the 3PL features of the verb. At the CP phase, the [P]

<sup>6</sup> We assume partitive as the default objective Case (Herschensohn 1996), a slight modification of Belletti (1988). The c-selectional features of the verb check the case of the verbal complement.

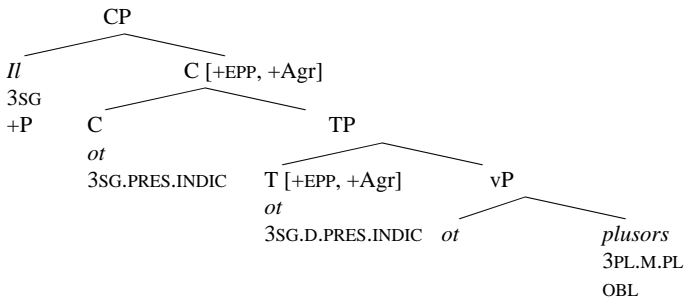
feature of *il* deletes the [EPP] feature of C, while the verb's interpretable mood checks off the [+Agr] feature.

(22)



In (23) the associate *plusors* is checked at the vP phase to eliminate uninterpretable oblique Case. At the TP phase, the interpretable tense feature of *ot* deletes [+Agr] of T, while the rich person features 3SG delete the [+EPP]. At the CP phase, *il* merges with C+T+v to delete uninterpretable [+EPP], while the mood feature deletes [+Agr] of C.

(23)



Our analysis of impersonal *il* and *ce/ço* as CP expletives shares the spirit of Piatt (1898), who notes that although *il* is not found in the very earliest French texts (such as the *Serments de Strasbourg*), the expletive *ce/ço* is already introduced, and by the 11th century, *ço* and *il* co-occur as neuters: “What has been given is sufficient to show that impersonal *il* was uniformly felt as a neuter, from its first occurrence to the time when the inflectional distinction between masculine and neuter began to be lost.... Such being the case, it could

not possibly have been the masculine pronoun put to impersonal uses” (Piatt 1898:63).

Since the presence of *il* in expletive constructions is only due to the [+EPP] feature in C (e.g., to maintain V2 word order in Arteaga 1994’s terms), we predict that *il* should not be expressed if another XP has fronted to CP. This prediction is generally borne out, for impersonal *il* in matrix clauses is usually unexpressed if not found in CP in OF.<sup>7</sup> A notable exception is the existence of matrix clauses (24) with encliticized subject pronouns in C (Barbara Vance, personal communication, 16 June 2003):<sup>8</sup>

(24) Encliticized subjects

- a. *Si ot il assez en la place barons et chevaliers qui*  
 thus had<sub>3SG</sub> EX many in the place barons<sub>M,PL,OBL</sub> and knights<sub>M,PL,OBL</sub> who  
*la voldrent retenir*  
 her<sub>F,SG,OBL</sub> wanted<sub>3PL</sub> to-retain  
 “Thus there were many barons and knights in the place who wanted to retain her.”  
 (*Queste* 13, 15; Vance 1997:246)
- b. *et a issir t’ en covient il o a morir*  
 and to go-out you<sub>2SG,OBL</sub> from-here is-necessary<sub>3SG</sub> it or to die  
 “and it is necessary for you to escape from here or else to die.”  
 (*Queste* 106, 16; Vance 1997:234)

The examples like those in (24) all have impersonal verbs constructed with XP (often an inversion trigger) in Spec CP and V+nominative pronoun in C. Noting that the interpretable person features of verb and pronoun match, Roberts (1993, extending Vance’s earlier analysis) proposes that these subject pronouns are cliticized from Spec IP. Roberts and Vance provide compelling evidence for the nominative nature of this impersonal *il*, which seems to us to

<sup>7</sup> Expletive *il* generally appears in subject position in embedded clauses. Jensen (1990:280) provides the following example:

(i) *qu’ il i avoit lettres entaillees*  
 that EX there were<sub>3SG</sub> letters<sub>F,PL,OBL</sub> carved<sub>F,PL,OBL</sub>  
 “that there were letters carved on it [the rock]”

Although this is outside the scope of the present paper, it appears to be part of a larger asymmetry found between expression and nonexpression of the subject pronoun in OF main clauses versus subordinate clauses, accounted for within GB theory by Adams (1987), Vance (1997), and Roberts (1993) by the inability of the verb to front to C° and license pro in subordinate clauses.

<sup>8</sup> Vance (1997:234) notes that postverbal pronominal subjects are “found only when the initial constituent refers directly back (or occasionally forward) to other elements of the discourse.”

be clearly distinct from expletive *il* of Spec CP. The role of these enclitic pronouns is beyond the scope of our paper.<sup>9</sup>

In our view, although the insights of Roberts (1993) and Arteaga (1994) are essentially correct, our analysis within the Minimalist Program provides a more principled means for correlating the diachronic loss of V2 and null subjects, since they are initially related to the rich verbal and nominal morphology of OF, while their loss is due to the erosion of that rich inflection. The MP is motivated by morphological features, interpretable and uninterpretable, so the richness of overt morphology will determine the options for distribution of uninterpretable features (the cause of displacement). This is a very significant insight because it derives the syntax from the morphology of the language, and it follows that a language like OF that has many interpretable features will in principle allow more uninterpretable ones than a morphologically impoverished language like MF. Finally, within the MP framework, the EPP is motivated by its connection to nominal Case and verbal inflection, both of which can satisfy it. For these reasons we believe that our analysis is superior to that of Roberts (1993) and Arteaga (1994) in the GB framework, in which movement possibilities and the EPP had to be stipulated.

#### 4. *Conclusion*

In this paper we have argued that the distribution of OF expletives and agreement with associates are a function of the parametric settings of T and C. The loss of nominal and verbal inflection led to the loss of V2 and null subjects in MF. Our analysis is preferable to earlier treatments in the GB framework because it motivates the uninterpretable features of Case, EPP, and Agr, and it relates them to their interpretable counterparts. We have argued that the richness of overt morphology determines the availability of uninterpretable features, and thus we derive the syntax directly from the morphology of the language.

---

<sup>9</sup> Although the evolution from OF to MF is outside the scope of this paper, it is interesting to note that this type of inversion is first seen with existential neuter constructions, as opposed to true impersonal verbs. We therefore find very appealing Barbara Vance's suggestion (personal communication, 26 June 2003) that these kinds of examples demonstrate that OF has already, in 13<sup>th</sup>-century prose, moved far enough toward requiring subject pronouns overall that the early OF system (with CP expletives) is no longer in effect, at least within existential constructions.

## REFERENCES

- Adams, Marianne. 1987. "From OF to the Theory of Pro-drop". *Natural Language and Linguistic Theory* 5.1-32.
- Alexiadou, Artemis & Elena Anagnostopoulou. 1998. "Parameterizing AGR: Word order, V-movement and EPP-checking". *Natural Language and Linguistic Theory* 16.491-539.
- , 2001. "The Subject in Situ Generalization and the Role of Case in Driving Computations". *Linguistic Inquiry* 32.193-231.
- Arteaga, Deborah. 1994. "Impersonal Constructions in Old French". *Issues and Theory in Romance Linguistics* ed. by Michael Mazzolla, 141-156. Washington, DC: Georgetown University Press.
- Arteaga, Deborah & Julia Herschensohn. 2001. "Case, Agreement, and Expletives: A parametric difference in French and English". Paper presented at the annual meeting of the Modern Language Association, New Orleans, December 2001.
- , 2003. "Diachronic shift of [+EPP] from Latin to Modern French". Paper presented at the West Coast Conference on Formal Linguistics, Tucson, September 2003.
- Belletti, Adriana. 1988. "The Case of Unaccusatives". *Linguistic Inquiry* 19.1-34.
- Boeckx, Cedric & Sandra Stjepanovic. 2001. "Head-ing Toward PF". *Linguistic Inquiry* 32.345-355.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.
- , 2000. "Minimalist Inquiries: The framework". *Step by Step: Essays on minimalist syntax in honor of Howard Lasnik* ed. by Robert Martin, David Michaels, & Juan Uriagereka, 89-155. Cambridge, MA: MIT Press.
- , 2001a. "Beyond Explanatory Adequacy". *MIT Working Papers in Linguistics* 20.1-28.
- , 2001b. "Derivation by Phase". *Ken Hale: A life in language* ed. by M. Kenstowicz, 1-52. Cambridge, MA: MIT Press.
- De Lage, Guy Raynaud. 1975. *Introduction à l'ancien français*. Paris: Société d'édition d'enseignement supérieur.
- Einhorn, E. 1974. *Old French: A concise handbook*. Cambridge: Cambridge University Press.
- Epstein, Samuel & Norbert Hornstein, eds. 1999. *Working Minimalism*. Cambridge, MA: MIT Press.
- Foulet, Lucien. [1919] 1995. *Petite syntaxe de l'ancien français*. Reprint, Paris: Honoré Champion.
- Groat, Erich. 1999. "Raising the Case of Expletives". *Working Minimalism* ed. by Samuel Epstein & Norbert Hornstein, 27-43. Cambridge, MA: MIT Press.
- Herschensohn, Julia. 1996. *Case Suspension and Binary Complement Structure in French*. Amsterdam: John Benjamins.
- Holmberg, Anders. 2000. "Scandinavian Stylistic Fronting: How any category can become an expletive". *Linguistic Inquiry* 31.445-483.
- , 2002. "Expletives and Agreement in Scandinavian Passives". *Journal of Comparative Germanic Linguistics* 4.85-128.
- Horning, Adolf. 1880. "Le pronom neutre *il* en langue d'oïl". *Romanische Studien* 4.229-272.
- Jensen, Frede. 1990. *Old French and Comparative Gallo-Romance Syntax*. Tübingen: Niemeyer.
- Koenenman, Olaf & Ad Neeleman. 2001. "Predication, Verb Movement, and the Distribution of Expletives". *Lingua* 111.189-233

- Martin, Roger. 1999. "Case, the Extended Projection Principle, and Minimalism". *Working Minimalism* ed. by Samuel Epstein & Norbert Hornstein, 1-25. Cambridge, MA: MIT Press.
- Moignet, Guy. 1976. *Grammaire de l'ancien français*. Paris: Klincksieck.
- Ndayiragije, Juvenal. 1999. "Checking Economy". *Linguistic Inquiry* 30.399-444.
- Piatt, Herman. 1898. *Neuter il in Old French*. Ph.D. Dissertation, University of Strasbourg.
- Roberts, Ian. 1993. *Verbs and Diachronic Syntax: A comparative history of English and French*. Norwell: Kluwer Academic Press.
- Togeby, Knud. 1974. *Précis historique de grammaire française*. Odense: Akademisk Forlag.
- Vance, Barbara. 1997. *Syntactic Change in Medieval French: Verb second and null subjects*. Boston: Kluwer.





# PARADIGMATIC AND SYNTAGMATIC RELATIONS IN ITALIAN VERBAL INFLECTION\*

LUIGI BURZIO  
*Johns Hopkins University*

## 0. *Introduction*

The shape of verbal stems can be determined in different ways, a representative spectrum of which is given in (1).

(1) Determining the shape of stems	Morphology	Phonology
a. GO, PAST ⇒ went	Suppletive	
b. /permit/-t ⇒ permi[r]-ed	Regular	✓ ± PU
c. /keep/-t ⇒ kep-t	Weakly suppletive	✓ + Residue
PU: /beep/-t ⇒ beep-t (Paradigm uniformity in force in large conjugations) Residue: <i>keep</i> , ∈ separate conjugation $C_x$		

The case of *went* in (1a) illustrates suppletive morphology. Here, the information PAST single-handedly determines the form of the verb, phonology playing no particular role. The case of *permitted* in (1b) illustrates regular morphology. Here, the information PAST determines the basic form of the affix, with the phonology plausibly handling the rest. We could say that the morphology supplies a /-t/, and that the phonology then imposes epenthesis as well as voicing, yielding [ɪd]. In American English, the phonology is also responsible for the flapping of the stem-final /t/ to [ɾ]. This allophonic process notwithstanding, with such morphologically regular items, the bulk of the phonology is typically inhibited by ‘paradigm uniformity’ (PU) effects (Burzio 2002a) that level the form of the stem. In further contrast, the case in (1c) can be regarded as ‘weakly’ suppletive. On the one hand, much of the information required is fully general: Morphologically, the affix can be taken to be the same as that in (1b), and the stem alteration can be straightforwardly attributed

---

\* Versions of this work were presented at the 33<sup>rd</sup> Linguistic Symposium on Romance Languages, Indiana University, April 2003, and at the 30<sup>th</sup> Incontro di Grammatica Generativa, University of Venice, Italy, February 2004. I am grateful to both audiences for constructive comments. Further improvements in both form and substance were made possible by suggestions kindly conveyed by the editors, and important clarifications owe to comments graciously provided by Vito Pirrelli.

to the phonology: closed syllable shortening, induced by the affix. On the other hand, however, there is a residue of lexical information that needs to be provided, so as to distinguish this special case from regular *beep*, which yields *beep*[t], and not *\*bept*, in violation of closed syllable shortening. This residue can simply be a specification that *keep* is a member of a special minor conjugation  $C_x$ , as indicated in (1). This will make it exempt from the PU effects that, as I argue later, rule over large conjugations, and are responsible for blocking *\*bept*.

In this chapter, I consider Italian verbal inflection in the three minor conjugations in *-ĕre*, *-ĕre*, *-ĭre* and examine the claim of Pirrelli and Battista (2000) (henceforth ‘P&B’) that stem shapes in these conjugations are determined purely as a function of ‘paradigmatic’ information, so that, for instance, the present indicative of a verb like *dolere* ‘ail’ in (2) would simply have four different suppletive forms of the stem, S1-S4, assigned to different cells in the paradigm as indicated. Throughout, S1—henceforth given in shaded cells for visual ease—will be the stem that also appears in the infinitive (my ‘S1’ is simply ‘S’ in P&B’s original text).

(2) P&B: A ‘paradigmatic’ account

Person	1	2	3	4	5	6
Affix	-o	-i	-e	-iámo	-éte	-ono
<i>dolere</i> ‘ail’	dolg	dwol	doł	dol	dolg	
4 stems	S2	S3	S4	S1	S2	

On P&B’s account, the form of the affixes in (2) would be incidental, and would play no (synchronic) role in determining the form of the stem. Reliance on stem indices as in (2) yields paradigmatic effects: Cells bearing the same stem index will host identical stems whatever the verb. In contrast to this view, I will argue that stem shapes in these cases are determined largely syntagmatically, with affixes playing a crucial role much as in the case of *kep-t* of (1c). At the same time, I will also partly concur with P&B, in recognizing that PU effects, and hence paradigmatic relations, are also at work. Crucial to my analysis is the assumption that lexicon and grammar do not partition the set of observed phenomena disjunctively. That is, the fact that some phenomenon may have a lexical component to it is not taken here to entail that it must be purely lexical. Rather, it is logically possible, and will be argued to be the case, that the grammar—phonology or morphology—can work alongside the lexicon. A special case of this was seen in (1c). The challenge taken up later is identifying a framework in which the attested interactions can be expressed.

The rest of the chapter is organized as follows. Section 1 introduces the analytical framework. Section 2 argues for the crucial role of affixes in determining the form of stems in Italian verbal inflection. Section 3 addresses the necessity for paradigmatic relations in addition to syntagmatic ones, arguing that these instantiate types of ‘Output-Output Faithfulness’ in Optimality Theory (OT). Section 4 addresses the total absence of stem alternations in the much larger *-are* conjugation and provides an account of this by which larger class size results in higher ranked Faithfulness, yielding total PU. Section 5 compares Faithfulness constraints with P&B’s coindexing of cells, showing that only the former is consistent with cases in which identity is only partial. Section 6 concludes the chapter.

### 1. *OT Constraints and entailments*

The analysis I present here relies on the framework of Burzio (2002a, 2002b, to appear), which is an extension of OT (Prince & Smolensky 1993) in a direction that would bring it closer to the architecture of neural nets, to which OT is independently related (Prince & Smolensky 1997). The central tenet of that approach is the hypothesis in (3).

(3) Representational Entailments Hypothesis (REH):

Representations are sets of entailments. Any representation AB corresponds to the entailments  $A \Rightarrow B$ ,  $B \Rightarrow A$ .

While a full justification of (3) is beyond the present goals, one critical piece of evidence for it is the fact that the effects attributed to Faithfulness constraints (Input-Output or Output-Output) do not seem definable in terms of a fixed rank, but rather have the properties of attraction between celestial bodies or magnets: the closer the two representations being compared, the stronger the effect. One example of this is the well-known syndrome called ‘Nonderived Environment Blocking.’ For instance, in Campidanian Sardinian an input /p/ both voices and spirantizes, yielding [β], while an input /b/ fails to spirantize (Burzio 2002b, citing Łubowicz 1999). That is, spirantization of /b/ ‘blocks’ unless the latter /b/ is ‘derived’ (from /p/). The ‘attraction’ characterization of this is that a candidate output with a [b] in it is under stronger attraction by its input if the latter input features the same /b/, but under a weaker attraction if the input has a /p/ instead—the ‘distance’ that weakens the attraction. Unlike spirantization, which can be blocked, voicing is evidently able to overcome even the maximal attraction. The REH (3) can characterize attraction effects as entailment summation. For instance, for a representation  $R1 = A, B, C$ , negation of C in another representation  $R2 = A, B, \neg C$  violates two of the

entailments generated by R1:  $A \Rightarrow C, B \Rightarrow C$ . In contrast, if R2 =  $A, \neg B, \neg C$  (R2 more distant from R1), then only one entailment will be violated by  $\neg C$ :  $A \Rightarrow C$ . The formerly violated  $B \Rightarrow C$  is now satisfied, since B no longer holds. In the Sardinian example, components B and C would stand for voicing and continuancy (spirantization) respectively, and A for the rest of the representation.

The hypothesis in (3) blurs the traditional distinction between representations or lexicon, and the grammar. The reason is that representations are, on this view, clusters of atomic constraints: the entailments, and hence also a form of grammar. The neural net affinity in (3) is in the fact that entailments and their ranks are analogous to weighted connections between units (if unit A is active, then a connected unit B must also be active). Also, in order for entailments generated by one representation to be at all relevant to other representations in the ways just described, it must be the case that the representations in fact share the same units (A, B, etc.), namely, that they are ‘distributed’ over those units, as in neural nets.

In addition to reconstructing the notion of Faithfulness in OT (both IO- and OO-) and accounting for the modulating effect of distance on their rank, the REH in (3) can also eliminate a rather obvious redundancy between OO-Faithfulness/PU constraints proposed in OT, which recapitulate essential facts about morphology, and the morphology itself, in the manner illustrated in (4).

(4) OO-FAITH/PU = Morphology

Representations	Entailments	Phenomena: Stem selection and PU
a. beep beep-ed	$Af_1 \Rightarrow / X \_\_\_$ $X = \text{Lexical category}$	violated by <i>keep/kep-t</i>
b. prend-o “I take” prend-i “you take”	$Af_1 \Rightarrow / X \_\_\_$ $X = / \_\_\_ Af_2$	violated by <i>veng-o/vien-i</i> “I come/you come”

Under the REH (3), pairs of representations like *beep/beep-ed* in (4a) will jointly generate entailments of the form: ‘A certain affix  $Af_1$  entails the presence to its left of a form  $X$  such that  $X$  occurs separately as a specific lexical category (here a verb) in the lexicon’ (cf. Burzio 2002b for details). Such entailments constitute simultaneously both subcategorization frames for the affix (traditionally part of morphology) and Output-Output Faithfulness (OO-FAITH) constraints in OT, requiring that the affixed stem must equal the lexical verb: a type of PU. Such entailments/constraints are violated by cases like *keep/kept* as stated in (4a). Turning to (4b), pairs of representations like Italian *prend-o/prend-i* “I take/you<sub>SG</sub> take” will similarly generate entailments

of the form ‘A certain affix  $Af_1$  entails the presence to its left of a form  $X$  such that  $X$  also occurs in the lexicon followed by a certain other affix  $Af_2$ .’ Such entailments are again simultaneously statements describing the workings of the morphology and OO-FAITH constraints in OT. Here, they impose uniform exponence of a stem in an inflectional paradigm, and are violated by any alternating stem such as that of *veng-o/vien-i* “I come/you<sub>SG</sub> come.” The REH (3) thus enables representations to condition other representations directly, without the intervention of an extrinsically defined ‘grammar.’ The ‘grammatical’ effect is the result of summation of identical entailments across the lexicon. Hence many regularities result on this view from the fact that the fundamental property of representational entailments is to penalize variation.

At the same time as they characterize the pressure for stem leveling or PU, entailments also characterize the pressure for affix leveling or syncretism, as shown in (5).

(5) Affixal syncretism  $\approx$  PU

Representation	Entailments	Phenomenon: Syncretism
<i>prend-o</i> “I take”	$X \Rightarrow Af_1$ * $X-Af_2$	violated by <i>prend-i</i> “you take”

In (5), a representation like Italian *prend-o* will generate an entailment of the form: ‘A certain specific form  $X$  must be followed by a specific affix  $Af_1$ .’ This entailment will then be violated by a form like *prend-i*, in which the same stem is followed by a different affix  $Af_2$ . This is taken to be the source of affixal syncretism, as occurs for example in the Italian singular present subjunctive *-a/-a/-a* compared with its indicative counterpart *-o/-i/-e*. There is no precedent for an account of syncretism within OT so far as I know, the most popular accounts to date being in terms of the ‘impoverishment’ rules of Distributed Morphology (Halle & Marantz 1993 and refs.), or the rules of ‘referral’ of Stump (2001 and refs.). On the reasons why syncretism occurs more in ‘marked’ categories (like subjunctive) than in unmarked ones (indicative), see Burzio (2002a, 2002b, to appear) and Tantalou and Burzio (2003), where this generalization is also derived from the REH. In essence, more marked categories are higher dimensional—they have a larger number of semantic components to them, producing larger numbers of entailments, and hence greater resistance to variation.

The structure of entailments can also derive Pānini’s principle, according to which specific information prevails over more general information, and thus in turn characterize morphological irregularity/suppletion—a form of specificity. This is illustrated in (6) and (7).

## (6) Suppletion by Pānini

Representations	Entailments	Phenomenon: Suppletion
went (GO, PAST) -t (PAST)	(GO, PAST) $\Rightarrow$ went (PAST) $\Rightarrow$ -t	Specific $\gg$ General (Pānini)

Any formal account of suppletion requires that a statement like ‘The past tense of GO is expressed by *went*’ prevail over the corresponding more general statement ‘The past tense of a verb is expressed by *-t* (or *-ed*),’ which would produce \**goed*. The former statement is more specific than the latter because it makes reference to the specific verb GO, rather than just the category ‘verb.’ The inherent edge enjoyed by specificity is referred to as Pānini’s principle. The REH (3) can substantiate Pānini’s principle in the fact that more specific information corresponds by definition to higher dimensional representations, producing a larger number of entailments. For instance, the representation *went* with its correlated semantics ‘GO, PAST’ is higher dimensional than the representation *-ed* with its semantics ‘PAST’ because it contains the additional component GO—an abbreviation for a complex semantic representation. Hence each of the two competing entailments in (6) effectively stands for a cluster of entailments of different dimensionality. Higher dimensionality provides a competitive advantage as illustrated in (7).

## (7) Deriving Pānini

	Representations	Entailments
a.	A, B, C, D, E	B $\Rightarrow$ A; C $\Rightarrow$ A; D $\Rightarrow$ A; E $\Rightarrow$ A
b.	$\neg$ A, B	B $\Rightarrow$ $\neg$ A

In the five-dimensional representation in (7a), component A is entailed by four other components, and hence four times, whereas in the two-dimensional representation in (7b), its rival  $\neg$ A would be entailed only once. Assuming that optimization works by minimizing the number of entailments violated, A will be optimal in (7a) despite violating the one entailment of (7b). We can deploy the schema in (7) over (6) by taking A and  $\neg$ A to represent *went* and *-t* respectively, B to represent the feature PAST, and CDE to represent the complex meaning GO. Other kinds of morphological irregularity, like the ‘weak suppletion’ of *kept* (1c), will also be characterizable in these general terms. A preliminary account of morphological irregularity along these lines was given in Burzio (2000), where it was argued that morphological irregularity can be attributed to Input-Output Faithfulness (IO-FAITH) dominating over Output-Output Faithfulness (OO-FAITH). This would result in the output [went] being faithful to the input /went/ rather than to other outputs

in which the past tense is expressed as [...ed]. The present approach is a follow-up on that earlier analysis, further reducing both types of faithfulness, IO-FAITH and OO-FAITH, to the elementary notion of entailment.

The discussion of morphological irregularity/suppletion in the preceding paragraph is obviously incomplete as it stands, as it provides no leeway for the regular or general to be asserted over the irregular or specific, failing to account for the fact that cases like *went*, and hence the effects of Pānini’s principle (Halle & Marantz 1993), are relatively confined. The needed (anti-Pānini) corrective is in the fact that regularities or generalization are by definition multiply instantiated, resulting in the more realistic competition in (8).

(8) Constraining Pānini

	Representations	Entailments
a.	$n * AB$	$n * (B \Rightarrow A)$
b.	$\neg A, B, C, D$	$B \Rightarrow \neg A; C \Rightarrow \neg A; D \Rightarrow \neg A$

Given a generalization  $AB$  instantiated  $n$  times, the number of entailments favoring  $A$  will be  $n$ , as in (8a), with the potential for a numerical override of the entailments asserting its rival  $\neg A$  from the representation in (8b). The latter could of course also be instantiated repeatedly, over time. Hence a full resolution of the competition will require quantifying frequencies of instantiation of both the generalization in (8a) and its potential violator (8b). These are known as ‘type’ and ‘token’ frequency, respectively. While this task is much beyond the present goals, these considerations nonetheless lay the basic groundwork, and will be sufficient for present purposes to express the competition between the morphologically regular—usually interpreted as a manifestation of some ‘grammar’—and the morphologically irregular—usually interpreted as a manifestation of the ‘lexicon.’ On the present view, the competition is unsurprising, given that entailments are the common currency for both lexicon and grammar.

An example of how entailments can bridge the divide between grammar and lexicon can be given by considering certain pre-affixal inserts that one finds in the Italian minor conjugations, specifically the inserts *-g-* and *-i-* of (9a) and (9b) respectively.

(9) Competition of pre-affixal inserts *-g-*, *-i-*:

- a. val-g-o ... val-g-ono “I am valued ... they are valued”  
(*salire, divellere, svellere, cogliere, togliere, scegliere, venire, rimanere, tenere, spegnere*)
- b. vol-i-o ... vol-i-ono  $\Rightarrow$  vollo (voglio) ... vollono (vogliono) “I want ... they want”  
(*piacere, tacere, giacere, solere, dovere*)



Such inserts occur in the present indicative in persons 1 and 6 with the verbs listed in (9a, b) respectively and perhaps a few others. The insert *-i-* is not observable directly, but rather only via the palatalization and/or gemination effects it produces. The process whereby the context *iV* produces palatalization/gemination of the preceding consonant is the one insightfully analyzed in Schein and Steriade (1986). In the present framework, the entailments generated by individual representations are assumed to undergo summation when they are identical, giving rise to higher order entailments, like those that were listed in (4) and (5). In the case of (9a-b) such entailments would have roughly the forms in (10a-b), respectively.

(10) Entailments as ‘grammar’:

- a.  $/l/n/\tilde{n}- \_ -o \Rightarrow g$
- b.  $/l/k/v- \_ -o \Rightarrow i$

Each of the statements in (10) reads: ‘The environment on the left of the arrow entails in it the presence of the element on the right of the arrow.’ The environment in (10a) is clearly amenable to simplification in terms of distinctive features (sonorant coronal other than *r*), whereas the one in (10b) is a bit more heterogeneous. The two statements in (10) are each equivalent to some piece of grammatical machinery—some kind of ‘readjustment’ rule. However, by themselves, such statements are insufficient to fully predict the outcome, because of both exceptions (cf. expected *piak-i-o*  $\Rightarrow$  *piáččo* ‘I please’ vs. unexpected *konduk-o*, not *\*konduččo* ‘I conduct’) and mutual conflicts. In particular, stems ending in *l* are targeted by both (10a) and (10b). This will require the cooperation of entailments that are specific to individual representations, such as the ones in (11), which draw the needed distinction between *val-ere* and *vol-ere*—a purely ‘lexical’ distinction.

(11) Entailments as ‘lexicon’:

- a.  $val- \_ -o \Rightarrow g$
- b.  $vol- \_ -o \Rightarrow i$

Note that while the item-specific statements in (11) may now seem sufficient, they in fact are not, since they fail to express the fact that the stems that select for these inserts tend to end in certain specific consonants. The statements in (10) are thus also needed alongside those in (11) to express the relevant generalizations. Both types of statements result from the entailments generated by specific representations and their summation. An overall account will feature further summation, with the entailments of (11) being added to those of (10) and thus breaking the tie in either direction: (11a) or (11b). In addition to

this, however, there are also full ‘Pāinian’ overrides of the entailments in (10): the pattern ‘...duk-’ being a general exception to (10b).

In summary, we find that morphological phenomena often have a hybrid character: on the one hand not totally random, and hence not amenable to a purely lexical analysis, while at the same time not totally regular, and hence not suitable for a purely grammatical analysis either. The case of English *keep/kept* in (1c) and that of Italian presuffixal inserts *-g-*, *-i-* are examples of such hybrid phenomena—what I have referred to as ‘weak suppletion.’ The frequent way in which lexicon and grammar thus appear to be finely interspersed suggests that, contrary to mainstream generative tradition, these two concepts do not correspond to discrete components of the language faculty, but are rather carved out of the same set of primitives. The REH (3) supplies those primitives in the form of the representational entailments.

## 2. *The syntagmatic role of affixes*

In this section I defend the relatively traditional view that affixes play a crucial role in conditioning the form of their stems against P&B’s claim that the stem form is purely a function of the paradigm cell in which it appears.

Interestingly, in the introductory part of their work, P&B give an assessment that seems much in line with the present one:

- (12) “The two dimensions [paradigmatic, and syntagmatic] turn out to be interlocked in a complex way, to define a *grammatical continuum*....” (P&B:307).

Given that P&B use the notion of ‘paradigmatic’ to refer to a type of suppletion as I indicated in connection with (2), while the notion ‘syntagmatic’ refers to a grammatical type of interaction, (12) seems a close restatement of my own conclusions in the previous section. In formulating their actual analysis, however, P&B seem to take a different position, suggesting a sharp separation between lexicon and grammar:

- (13) “All alternating stem roots which are not accountable in terms of exceptionless phonological rules of Italian are to be considered as independent **B[asic] S[tems]** in Aronoff’s sense.” (P&B:323)

Unlike the statement in (12), the one in (13) seems to exclude lexical-phonological hybrids. Other work by the same authors reveals that there is in fact no contradiction between the two statements. In particular, Pirrelli (2002) outlines a conception similar in spirit to the present one in seeing a continuum of possibilities, but crucially different on the particulars of what defines the continuum. In that conception, there is a competition between syntagmatic

factors, which presumably define calculations from a unique input to a unique output, for example, /ven-iamo/  $\Rightarrow$  [veniamo], and syntagmatic factors, which define both identity of forms: [veng-o] = [veng-ono] (same stem) and differences: [ven-iamo]  $\neq$  [veng-o] (different stems). This characterization is the one echoed in P&B's statement in (12). In the present system, however, the competition is not merely between syntagmatic and paradigmatic factors, but rather three-way: IO-FAITH (lexicon); MARKEDNESS (phonology); OO-FAITH (morphology). In P&B's conception, the paradigmatic dimension seems closely tied to the notion of lexicon. In contrast, in the present one, paradigmatic relations are rather subsumed under morphology: part of the grammar, though emergent from the lexicon via entailment summation. The multiplicity of correspondence among paradigm cells given by P&B's coindexing follows here from transitivity of the notion of identity. If, as a result of the entailments,  $X$  must equal  $Y$  in  $X-Af_1/ Y-Af_2$  (no stem allomorphy), and if at the same time  $X$  must also equal  $Z$  of some  $Z-Af_3$ , then  $Y$  must also equal  $Z$  by transitivity. This is the source of 'multiple correspondence' in the sense of Burzio (1998), a by-product of a morphology based on the OT notion of (OO-)Faithfulness, though not of a morphology that was based on traditional rewrite operations. Hence the statement in (12) parallels the present position that there is a morphology-lexicon competition with a continuum of outcomes, while the statement in (13) does not parallel the present position that there is comparable competition between phonology and the rest: morphology or lexicon. But there is no contradiction between the two statements. Rather, it is just that the phonology is not factored into P&B's system of fine-grained interactions. Instead, it is regarded as an alien element whose character is already known in advance: It only works in an 'exceptionless' way, and herein lies the disagreement. In what follows I will argue that the position in (13) is not tenable empirically, as it misses crucial 'syntagmatic' regularities of the phonological kind. Before undertaking that task, however, I consider the P&B analysis in more detail.

The main feature of P&B's analysis is an attempt to characterize inter-paradigmatic relations in terms of rules that re-index cells. Their re-indexing schema for the present indicative is given in (14).

- (14) P&B's re-indexing schema
- |    |               |    |   |               |    |
|----|---------------|----|---|---------------|----|
| S3 | $\Rightarrow$ | S1 | Stipulation:<br><br>* S4 $\Rightarrow$ S3 |               |    |
| S3 | $\Rightarrow$ | S2 |   | $\Rightarrow$ | S1 |
| S4 | $\Rightarrow$ | S2 |   |               |    |
| S4 | $\Rightarrow$ | S1 |   |               |    |

The rules in (14) would make it possible to characterize simpler paradigms from more complex ones in the ways that we will see later. Each rule states that all the cells that are occupied by a stem  $S_n$  in the paradigm of some verb may, with some other verb, collapse with the cells occupied by stem  $S_m$ , where  $m$  is a smaller number than  $n$ . The input to the re-indexing rules in (14) is the four-stem paradigm structure in (2), from which simpler paradigm structures would be derivable by such re-indexing. P&B note, however, that the re-indexing ‘S4  $\Rightarrow$  S3’ (extend stem 3 to the cells formerly occupied by stem 4) is conspicuously unattested, a fact that for them requires the stipulation indicated in the right-hand box in (14).

I note first that while this approach may seem interesting, its explanatory power is in any event limited, since it is true by definition that a simpler paradigm is relatable to a more complex one by simplification. All such power would reside in the restricted format of the rules, in which all re-indexing is downward. Even this asset is questionable, given that the schema itself is the main reason for assigning specific indices to the various stems. Be that as it may, in the rest of the discussion I will rather focus on the fact that the schema in (14) is inadequate empirically, since it directly predicts nonexistent paradigms, in contrast to the syntagmatic approach that correctly excludes them.

I begin by considering the four-stem paradigm of (2), from which others would be derived in the P&B analysis. This paradigm is repeated in (15), with an outline of the syntagmatic analysis on the left of the table (shaded cells = stem of infinitive = S1).

(15) 4 stems =  
3 syntagmatic factors

Person	1	2	3	4	5	6
Affix	-o	-i	-e	-iámo	-éte	-ono
P&B's stems	S2	S3		S4	S1	S2
dol-ére "ail"	dolg	dwol		dołł	dol	dolg
a. Stem stress	Y	Y	Y			Y
b. -g- insert	Y					Y
c. Palatalization/gemination before iV				Y		

In a syntagmatic analysis, the fact that the paradigm in (15) attested with *dolere* consists of four subparadigms will require at least three differentiating factors, each splitting the paradigm in two. The actual factors are listed in (15a-c). One is stress (15a), and its role in the diphthongization  $o \Rightarrow wo$ . Consider here that the stem is stressed in persons 1, 2, 3, 6, in turn reflecting the fact that the affixes bear stress only in persons 4 and 5. In Italian, stress is

partly lexicalized and inflectional affixes take priority over stems in controlling it (Burzio 1998). However, diphthongization only occurs in persons 2 and 3 (P&B's S3), because (in Italian, as opposed to Spanish) an open syllable is also required in addition to stress, persons 1 and 6 having closed syllables. The second factor at work (15b) is the *-g-* insert discussed earlier, which, by creating the just noted closed syllables, thus distinguishes person 1, 6 from 2, 3, and hence P&B's S2 from S3. The third factor (15c) is palatalization and/or gemination before the sequence *iV*, also discussed earlier in connection with the *-i-* insert. Such an environment occurs only in person 4, where the affix is responsible for creating it, whence S4. Person 5 is unaffected by any of these factors, leaving the stem in the same (unstressed) environment as the infinitive, whence stem S1 in that person.

From the syntagmatic point of view just outlined, we will now expect simpler paradigms whenever one of the factors in (15) is inoperative or appropriately changed. This is true of each of the three-stem cases in (16)-(18), which are alternatively handled by P&B's re-indexing rules reported in each case.

(16)3 stems =

2 syntagmatic factors

Person	1	2	3	4	5	6
Affix	-o	-i	-e	-iámo	-éte	-ono
P&B's re-indexing: S3 ⇒ S2	S2		S4		S1	S2
dov-ére "have to"	dev		dobb		dov	dev
a. Stem stress	Y	Y	Y			Y
b. Palatalization/gemination before <i>iV</i>				Y		

The verb in (16) is not one of those that insert *-g-*, which were given in (9). With that factor removed, the separation between persons 1, 6 and the other persons with unstressed affixes (2, 3) will disappear, corresponding to removal of the separation between P&B's S2 and S3 of (15) as if the re-indexing rule in (16) had applied. The other two factors, listed in (16a-b), continue to play the same role as in (15), although their effects are now slightly different. Here stress controls the *e/o* variation: *e* labializes to *o* before labials, though not in stressed positions, which are notoriously more resilient to change. As for (16b), it is a general fact about labials that they do not palatalize where other consonants do. Hence only gemination occurs. However, fairly generally *v* only finds *bb* as its geminate counterpart in Italian (historically, singleton *b* but not geminate *bb* spirantized intervocalically). To be sure, some of these alternations are of limited productivity, a point correctly made by P&B, but in the present framework this does not disqualify them from having a syntagmatic

basis along with some lexical basis, as we saw in the discussion of the inserts of (9). That is, in the present context, lexical entailments can be recruited to assist an otherwise weak phonological effect, without making the alternation purely lexical, a point that I return to in section 4.

In the next case in (17), the *-g-* insert is reinstated, but the palatalization/gemination is suppressed, as the stem-final *n* does not participate in this process. Some lexical control, here inhibitory, may again be involved.

(17) 3 stems =  
2 syntagmatic factors

Person	1	2	3	4	5	6
Affix	-o	-i	-e	-iámo	-fte	-ono
P&B's re-indexing: S4 ⇒ S1	S2	S3		S1		S2
ven-íre "come"	veng	vyen		ven		veng
a. Stem stress	Y	Y	Y			Y
b. -g- insert	Y					Y

In (17), stem stress and *-g-* insert jointly control diphthongization (in this case, *e* ⇒ *ye*) much as in (15), but the effect that sets person 4 apart from 5, palatalization/gemination, is lacking, as /n/ is not affected. This is correctly describable by means of P&B's re-indexing rule in (17). A third logical possibility, consisting of only a *-g-* insert and palatalization/gemination without any stress-induced effects, appears not to be instantiated. This possibility would put persons 2, 3, and 5 together, and would be describable by P&B's rule 'S3 ⇒ S1' in (14). I take this to be an accidental gap. Since this discrepancy does not distinguish the two approaches, it will not affect the rest of the discussion.

Yet another logical possibility captured by both approaches is instantiated in (18).

(18) 3 stems =  
2 syntagmatic factors

Person	1	2	3	4	5	6
Affix	-o	-i	-e	-iámo	-íte	-ono
P&B's re-indexing: S4 ⇒ S2	S2	S3		S2	S1	S2
solere "to be used to"	solɰ	swol		solɰ	sol	solɰ
a. Stem stress	Y	Y	Y			Y
b. -i- insert + Palat./ Gemination before <i>iV</i>	Y			Y		Y

This case is exactly like the one in (15), except that here the insert is *-i-* instead of *-g-*. As a result of this, persons 1, 6, which receive the insert, become palatalizing/geminating environments just like person 4 in which the *i* is part of the affix, whence the same stem S2 in those three persons, as described by P&B's re-indexing. Because gemination creates closed syllables just like the

-g- insert, the distribution of diphthongization in S3 (18), is the same as that of (15). As in the latter case, S1 has no diphthong because it is unstressed.

Alongside the paradigms in (18) that are consistent with both approaches, however, the following three-stem paradigms do not exist as predicted only by the syntagmatic approach. The question marks in (19) convey the hypothetical nature of such verbs (pal/gem = palatalization/gemination).

Person	1	2	3	4	5	6
Affix	-o	-i	-e	-iámo	-éte/-íte	-ono
S2 ⇒ S1	S1	S3		S4	S1	
*?ol-ere/ -ire	?ol	?wol		?oĥ	?ol	
*S4 ⇒ S3	S2	S3			S1	S2
*??ol-ere/ -ire	??olg	??wol			??ol	??olg

(19) Nonexistent three-stem paradigms

- a. If 1, 6 = 5, then no insert, then 1, 6 = 2, 3
- b. If 2, 3 ≠ 5 (must be stress) and 2, 3 = 4 (no pal/gem), then 4 = 5

The nonexistent paradigm in (19a) is straightforwardly derivable from the P&B re-indexing rule given. Although this rule has not been utilized so far, we will see later that it cannot be eliminated from the schema in (14). This therefore constitutes an incorrect prediction. The syntagmatic account is straightforward, as given synoptically on the left-hand side in (19a): In order for the stem of persons 1 and 6 to equal that of person 5, there must not be an insert, but then persons 1, 6 cannot be separated from persons 2, 3 as in (19a) by any of the available factors: 1, 2, 3, 6 are all unstressed environments. We see later that one more potential factor needs to be recognized, involving palatalization before front vowels. Yet this factor will be of no help in (19a), since it will only place persons 2, 3 with 5 incorrectly (while 4 might differ by gemination). Hence paradigm (19a), allowed by P&B's re-indexing, is correctly excluded by the syntagmatic analysis.

The nonexistent paradigm in (19b) is banned by P&B's stipulation that the re-indexing rule needed is absent from the schema, as was shown in (14). But the syntagmatic approach is again superior as it requires no such stipulation: If the stems of persons 2, 3 differ from that of person 5, it can only be because stress is the relevant factor. Stress, however, cannot distinguish persons 4 and 5 as in (19b). The only such distinction can be drawn by palatalization/gemination, but that in turn cannot be involved, since it would also separate 2, 3 from 4 incorrectly. Hence (19b) is also correctly excluded. Note that palatalization before front vowels (20b') would provide again no recourse, since it would group 2, 3, and 5 together, also incorrectly.

We now turn to two-stem paradigms. In P&B's system, these require two re-indexing rules. In the syntagmatic approach, these will follow from the

presence of a single differentiating factor. I begin again with the cases that are correctly predicted by both systems, given in (20).

	Person	1	2	3	4	5	6
(20) Attested two-stem paradigms	Affix	-o	-i	-e	-iámo	-éte/-íte	-ono
a.	S3 ⇒ S2, S4 ⇒ S1	S2			S1		S2
Stress: 1, 2, 3, 6	sed-ére "sit"	syed			sed		syed
b.	S4 ⇒ S1, S3 ⇒ S1	S2	S1			S2	
-g- insert: 1, 6	val-ére "be worth"	valg	val			valg	
b'. Pal: 2, 3, 4, 5	krésc-ere "grow"	kresk	krešš			kresk	
c.	S4 ⇒ S2, S3 ⇒ S1	S2	S1	S2	S1	S2	
-i-; pal/gem: 1, 6, 4	piač-ére "please"	piačč	piač	piačč	piač	piačč	

While these follow from P&B's pairs of re-indexing rules given above each verb, the syntagmatic accounts are also straightforward as indicated in each case. Case (20a) is like (17) without the -g- insert, which produced closed syllables in persons 1, 6. Without it, diphthongization will now occur in all stressed stems: S2. Cases (20b) and (20b') have identical paradigms but for different reasons. In (20b), the -g- insert separates persons 1, 6: S2, from all others. In (20b') palatalization before front vowels separates all other persons: S1, from 1, 6. The gemination generally induced by the person 4 affix fails to produce a separation in this case, because simple palatalization of the cluster *sk* also yields geminate *šš*. In (20c), an -i- insert in persons 1, 6 produces the same geminating environment as in person 4 (as in (18)): S2. The other persons feature a singleton palatal before a front vowel, just like the infinitive: S1.

In contrast to those in (20), the following two-stem paradigms are unattested, however.

	Person	1	2	3	4	5	6
(21) Nonexistent two-stem paradigms	Affix	-o	-i	-e	-iámo	-éte/-íte	-ono
a. If 4≠5 (gem/pal) then 2, 3≠4	S4 ⇒ S2, S3 ⇒ S2	S2			S1	S2	
	*?ol-ere/ -ire	?oľ			?ol	?oľ	
b. If 1≠2, then either (insert) 6≠4, 5; or (pal) 2, 3 = 5	S4 ⇒ S1, S2 ⇒ S1	S1	S3		S1		
	*?inč-ere/ ire	?ink	?inč		?ink		

Such paradigms can be directly generated by P&B's pairs of rules given in each case. From the syntagmatic point of view, their nonexistence is again obvious. In the case of (21a), the needed distinction between persons 4 and 5 can only be produced by palatalization/gemination before *iV* in person 4. But the latter is never extendable to persons 2, 3 as required in (21a), only to person



1, 6 via the *-i-* insert (as in (20c) above). In the case of (21b), there are only two ways to produce the necessary separation of person 1 from 2, 3: either via an insert (persons 1, 6), in which case person 6 will also have to differ from 4, 5, incorrectly; or via palatalization before front vowels (as in (20b')), in which case persons 2, 3 will have to equal at least 4 (5 could geminate in addition), also incorrectly. Hence both paradigms are correctly excluded by the syntagmatic analysis.

In summary, P&B's re-indexing rules that take no account of the environment created by affixes generate several unattested present-indicative paradigms in addition to requiring the stipulation in (14). In contrast, a syntagmatic analysis along traditional lines that takes the form of stems to be derivable from the environment created by the affixes and the pre-affixal inserts will correctly exclude such paradigms and require no stipulation, while also accounting for all attested paradigms. We have seen that in many cases this requires that the lexicon play a role in promoting the specific phonological process, however, since the latter is not fully productive in the language.

An even stronger argument for the role of affixes than the one just given is provided by the observation that when affixes level, so do stems. This observation can be made by comparing the present indicative, which was reviewed earlier, with the present subjunctive, as in (22) (SP = subparadigm).

(22) Indicative versus  
subjunctive  
SP: {1, 6}, {2, 3}, {4}, {5}  
SP: {1, 2, 3, 6}, {4, 5}

- a. Stress: {1, 2, 3, 6}  
b. Pal/gem: {4, 5}  
c. -g- insert: {1, 2, 3, 6}

Person	1	2	3	4	5	6
Pres. ind.	-o	-i	-e	-iámo	-éte/-íte	-ono
Pres. subj.	-a	-a	-a	-iámo	-iáte	-ano
2 stems	S2		S1		S2	
sed-ére "sit"	sied		sed		sied	
krésc-ere "grow"	kresk		krešš		kresk	
ven-íre "come"	veng		ven		veng	

In (22), the top two rows compare the present indicative and present subjunctive affixal paradigms. The subjunctive is more leveled compared with the indicative—a case of the affixal syncretism that was briefly discussed earlier in (5). Consider in particular that the present-indicative affixes create a maximum of four different environments for the stem as follows: (a) stressed and followed by a back vowel as in persons 1, 6; (b) stressed and followed by a front vowel as in persons 2, 3; (c) unstressed and followed by *iV* as in person 4; (d) unstressed and followed by a simple front vowel as in person 5. From the present point of view this is the exact reason why the present indicative has a maximum of four different stems: the four subparadigms given on the left in

(22). The previous argument was that from P&B’s point of view this exact relationship between affixes and stems is a sheer coincidence. The subjunctive adds considerably to the degree of coincidence and hence to the strength of the argument, as I discuss next.

Unlike their indicative counterparts, the six subjunctive affixes create only two different environments for stems: (a) stressed and followed by a back vowel, as in persons 1, 2, 3, 6; and (b) unstressed and followed by *iV*, as in persons 4, 5. We can thus correctly predict that the present subjunctive will exhibit a maximum of two stems: the two subparadigms given on the left in (22). In particular, we predict that any of the three differentiating factors that were considered before—stress, palatalization/gemination, and *-g-* insert—will split the paradigm in exactly the same way, precisely as in (22a, b, c) respectively. Note that the environment for insertion of *-g-* and (as we see next) *-i-* of (10) needs to be generalized to a following back vowel from just [o], but this is straightforward. For the P&B analysis, these precise correlations are indeed inexplicable, since these are all just ‘basic stems’ (13), not cases of surface allomorphy.

Consider now also the leveling from two stems to just one in the subjunctive, illustrated in (23).

(23)Leveling the subjunctive

SP: {1, 2, 3, 6}, {4, 5}

a. *-i-* insert: {1, 2, 3, 6} = {4, 5}

b. None of the above

Person	1	2	3	4	5	6
Pres. subj.	-a	-a	-a	-iámo	-iáte	-ano
??S1 ⇒ S2	S2					
	piač-ére			piačč		
S2 ⇒ S1	S1					
	prénd-ere			prend		

There are two possibilities for leveling. One arises when *-i-* is inserted, as with the verb in (23a) (which was listed in (9b)). Such insertion, now occurring in persons 1, 2, 3, 6 because the affixes create the same environment for insertion, will make the environment of the latter persons identical to the environment of persons 4, 5, namely the environment *\_iV*, resulting in uniform palatalization/gemination throughout, as in (23a). In terms of re-indexing rules, the one that would be required here is one that would actually falsify P&B’s own schema if extended to the subjunctive. The reason is that the stem *piačč* of the subjunctive is not the one of the infinitive (*piač-ére*) and hence cannot be S1 on P&B’s criteria, but must rather be S2. Thus, in order for the simpler paradigm in (23a) to be derived from the more complex one of (23a-c) similarly to P&B’s derivation of indicative paradigms, re-indexing must go from S1 to S2, contrary to (14).

The second type of leveling, illustrated in (23b), occurs when (just as in the indicative), the verb is not prone to any of the effects of (22a-c)-(23a). The stem of the infinitive will then occur throughout. The re-indexing rule needed in this case is ‘S2  $\Rightarrow$  S1.’ This is in line with P&B’s schema in (14). Recall, however, that this was one of the rules that produced unattested paradigms in the indicative as in both (19a) and (21b). The effect in (23b) now indicates that such a rule cannot simply be dropped from the schema in (14), thus confirming its problematic nature.

In conclusion, paradigmatic relations are at least insufficient and syntagmatic relations are thus necessary to deal with stem alternations in the Italian minor conjugations. We have seen that such alternations are substantially predictable from the form of the affixes, modulo some lexical and other specific information like (11) and the correlation between singleton *v* and geminate *bb* of (16). In contrast, P&B’s ‘suppletive’ analysis enriched with the re-indexing schema in (14) both treats a complex set of stem-affix correlations as accidental and incorrectly generates a number of unattested paradigms.

### 3. *Rescuing paradigmatic relations*

The previous arguments notwithstanding, P&B are correct in pointing to the need for paradigmatic relations in at least two types of cases, in addition to a third one that will be discussed in section 5. The first case is illustrated in (24).

(24) Past participles have the same stem as their infinitives

Affix	Infinitive: <i>-ere/-ire</i>	Participle: <i>-uto</i>	Syntagmatic predictions
P&B	S1		
a.	krešš		*kresk-uto
b.	ven		*ven-g-uto

As indicated in (24), past participles do not follow the predictions of the strictly syntagmatic analysis, but rather have stems that are identical to that of the infinitive. Remember that in the case of *kréšš-ere* (*creocere*) ‘‘grow’’ (24a), palatalization is attributed to the presence of a front vowel: (22b), (20b’). The back vowel *u* of the participial affix should therefore yield \**kresk-uto*, incorrectly. Similarly, in the case of *ven-ire* ‘‘come’’ (24b), the back vowel of the participle should trigger insertion of *-g-*, as with the back vowels of (22c) or (17), also incorrectly. In the P&B framework, past participles can simply be assigned the same stem as the infinitive S1, as indicated in (24), affixes playing no role. In the present framework, along with syntagmatic relations,

paradigmatic ones can be expressed in terms of the PU of (4), a form of OO-Faithfulness. The participle of (24a) can then be analyzed as in (25):

(25)

Infinitive: [krešš- ], Input: /kresk-uto/	PALATALIZE	OO-FAITH (Infinitive)	IO-FAITH
a. kreskuto		*	
b. $\text{\textcircled{r}}$ kreššuto			*

In (25), the PALATALIZE constraint must dominate IO-FAITH to induce palatalization in the infinitive and the other cases discussed earlier. If OO-FAITH relating a participle to its infinitive also dominated IO-FAITH, then candidate (b) will be optimal as desired.<sup>1</sup> The question is of course why OO-FAITH should be higher ranked in the case of participle-infinitive pairs than in other cases, where stems vary freely. While a fully principled answer to this question must await further work, I note that there is independent evidence that participles are in strong correspondence with their infinitives, shown by paradigms like (26), analyzed in Burzio (1998, 2003).

- (26) a. v́inc-ere “win<sub>INF</sub>”  
 b. v́in-t-o “won<sub>PART</sub>”  
 c. v́inc-it-óre “winner”

Many verbs whose infinitive ends in unstressed *-ere* have syncopated participles like the one in (26), where the participial affix is *-t-* instead of regular *-út-*. In the work cited, I follow DiFabio (1990) in taking such syncopes to occur so as to allow the participle to be accentually faithful to the infinitive, avoiding the stress shift that nonsyncopated *-út-* would cause, while derivatives like (26c) are faithful to both (26a) and (26b) as I discuss later. This evidence

<sup>1</sup> The use of IO-FAITH as in (25) requires a brief comment given the fact that the present framework rejects the notion of ‘underlying representation’ (UR). Without UR, surface allomorphs can benefit from independent inputs. When allomorphy is suppletive, two surface allomorphs [A], [B] result from separate inputs /A/, /B/ under the ranking IO-FAITH >> OO-FAITH (cf. discussion of (6)). When allomorphy is not suppletive, two surface allomorphs [A], [A’], which are either identical or differ predictably, result from some input /A/ and the ranking OO-FAITH >> IO-FAITH. Discordant inputs /A/, /B/ are still entertainable as before (Richness of the Base), but their effects will be suppressed by dominant OO-FAITH. The issue of which surface allomorph, [A] or [A’], may receive the input /A/ under these circumstances may be indeterminate, but this is of no particular consequence. In this light, the IO-FAITH of (25) is then to be construed as abbreviating the joint effects of some input /kresk-/ “grow” of indeterminate affiliation over the paradigm, and a set of OO-FAITH relations to the surface allomorphs that instantiate that input. The gist of (25) is that the specific OO-FAITH to the stem of the infinitive outranks the more general effect just described as IO-FAITH.

is parallel to that of (24): The material preceding a participial affix must be identical to the material preceding the infinitival affix. However, unlike the correspondence of (24), that of (26a-b) cannot be expressed by P&B's coindexing, given the misidentity *vinc/vín(t)*. In contrast, such facts are consistent with OO-FAITH, assuming that undominated constraints on syllable structure rule out *\*vinc.to* for (26b). Further evidence favoring OT's faithfulness over indexing is considered in section 5. Also relevant is the fact that certain nominal derivatives like the one in (26c) take both the participle and the infinitive as their bases simultaneously—the 'multiple correspondence' alluded to earlier. The link to the participle is shown by the fact that the *-it-* of (26c) is an allomorph of the *-t-* of (26b), while the link to the infinitive is shown by the consonant *c* ([č]), not present in the participle. If (26c) is in correspondence with both (26a) and (26b), then (26b) will be in correspondence with (26a), by the transitivity inherent in the present system. P&B's coindexing will be ineffective in this case as well: The complex stem *vinc-it-* is in perfect correspondence neither with *vinc-* of (26a) nor with *vin-t-* of (26b)—it is only in partial correspondence with both.

The second case requiring paradigmatic relations involves a nonstandard variety of Italian (a working-class dialect, so far as I am aware) in which the present subjunctive affixes *-a*, *-ano* of (22) are replaced by *-i*, *-ino*, respectively. As P&B note, this switch does not affect the form of the stem, contrary to the strictly syntagmatic predictions, as indicated in (27).

(27) Nonstandard affixes leave stems unaffected:

Affix	Nonstandard <i>-i</i> ; <i>-ino</i> for <i>-a</i> ; <i>-ano</i>	Syntagmatic predictions:
P&B	S2, like persons 1, 6 of ind.	
a.	veng	*vyen-ino
b.	kresk	*krešš-ino

Syntagmatically, the front vowel of the affix in (27a-b) should trigger just the same behavior as the front vowels of the indicative, respectively: no insertion of *-g-*, with consequent diphthongization as in (17) *vyen-i*, hence *\*vyen-ino*; and palatalization as in (20b') *krešš-i*, hence *\*krešš-ino*. In contrast, in P&B's system, in which affixes play no role, the stem may remain the same S2 as in the standard dialect.

To make sense of this phenomenon from the present perspective, consider that the nonstandard subjunctive affixes instantiate a type of syncretism in which the affixes of the more numerous *-áre* conjugation are extended to the minor conjugations in *-ere*, *-ére*, *-íre*, as shown in (28).

(28) Conjugation

		Present-subjunctive affixes					
		1	2	3	4	5	6
a.	<i>-áre</i>	-i	-i	-i	-iámo	-iáte	-ino
b.	Minor, nonstandard	-i	-i	-i	-iámo	-iáte	-ino
c.	Minor, standard	-a	-a	-a	-iámo	-iáte	-ano

The transconjugational syncretism of (28a-b) is somewhat similar to the one already observable in persons 1-3 of each conjugation and discussed earlier. Like the former, it follows from entailment satisfaction, though space limitations preclude a full discussion of this here. Note as well that the standard affixes are themselves transconjugationally syncretic in various ways already: by being the same across the three minor conjugations (28c), and by being the same in persons 4 and 5 for all conjugations. The phenomenon of (27a) can be accounted for from this general point of view along the lines of (29), where ‘SYNC(RETIZE)’ encapsulates the factors that lead to syncretism: in practice imposing the affixes of the ‘major’ *-are* conjugation on verbs like *ven-ire*.

(29)

OO-FAITH: [veng-] (1, 6 ind.) IO-FAITH: /ven-ano/	SYNC: {-i, ...-ino}	OO-FAITH (1, 6 Indic.)	-g- ⇒ / __ [+back]	IO-FAITH
a. veng-ano	*			
b. vyen-ino		*		*
c. $\varnothing$ veng-ino			*	*

In (29), candidate (a) is the correct form for the standard dialect. It satisfies IO-FAITH since both stem and affix match the input (in the sense of fn. 1). It also satisfies the distributional requirement for the *-g-* insert: before back vowels only. In addition, it satisfies OO-FAITH of its stem to the stem of persons 1, 6 of the present indicative. More accurately, though, it contributes to the emergence of such a constraint, by the entailments it generates. In the standard variety, the identical stems in *veng-ono* (6, ind.), *veng-ano* (6, subj.) follow syntagmatically: the back vowel of both affixes. Identity relations, however they arise, generate entailments that make them self-sustaining by the schema in (4b) (and more detailed discussion in Burzio to appear). Candidate (a) thus only violates SYNC: a pressure to simplify the affixal system ultimately also akin to OO-FAITH. It wins in the standard variety, where SYNC is ranked below IO-FAITH, but loses in the nonstandard variety where that ranking is reversed. Candidate (b) complies with SYNC and shows regular application of *-g-* insertion: not before front vowels, and consequent regular diphthongization. It thus violates IO-FAITH for both stem and affix. It is excluded by the violation

of OO-FAITH as its stem does not correspond to [veng-] of persons 1, 6, present indicative. Candidate (c) thus wins despite the overapplication of *-g-* insertion and the unfaithful affix.

The analysis in (29) embodies the uncontroversial claim that speakers form a mental grammar that is maximally consistent with the data to which they are exposed. It also embodies the further claim, inherent in the REH (3), that whatever identity relations have a statistical presence in the data, also have, ipso facto, a grammatical status, expressible as faithfulness constraints in the OT formalism. The nonstandard variety in question constitutes a type of language change from the standard, which must therefore have provided the input data. The fact that the nonstandard maintains the same paradigmatic relations as the standard means that paradigmatic relations enter into the mental computation: P&B's point. In the present system, unlike in P&B's, the fact that syntagmatic relations do not obtain in this case (cf. inappropriate *-g-* insert) does not mean, however, that they do not exist, but rather only that they are outranked. The evidence for the present ontology rather than P&B's is that the syntagmatic relations violated in (29) are necessary elsewhere, as was shown earlier. Standard and nonstandard varieties must therefore differ by their constraint rankings. The account of *kresk-ino* (27b) would be fundamentally similar to that of *venghino* (27a) given in (29).

In sum, the present system has the resources to express both syntagmatic and paradigmatic relations. These relations are often in competition, the former demanding stem alternations, the latter aiming to prevent them. The present analysis has defended the claim that in the minor conjugations syntagmatic relations distinctly have the upper hand, paradigmatic effects arising only in some special cases. We see next that, in the *-are* conjugation, the tables are turned.

#### 4. *Why weak suppletion does not extend to large classes*

In contrast to what I have been calling the three 'minor' conjugations, the conjugations in *-are*, more numerous than the others by orders of magnitude, is totally leveled, featuring the same stem as the infinitive throughout, with very few exceptions. The P&B account describes this fact by postulating application of all re-indexing rules in this conjugation, resulting in the single S1. Obviously, this provides little explanation. The entailment-based approach, on the other hand, can capture the correlation between leveling and class size by way of the fact that entailment summation automatically assigns the appropriate statistical weight to grammatical effects. When class size reaches a

certain threshold, any alternation that has a lexical component will be swamped by grammatical pressures.

I illustrate this point with the case of palatalization/gemination before *iV*. In the minor conjugations, such an effect is under lexical control, as shown in (30a-a'), but in the *-are* conjugation, it is totally absent, as in (30b).

(30)

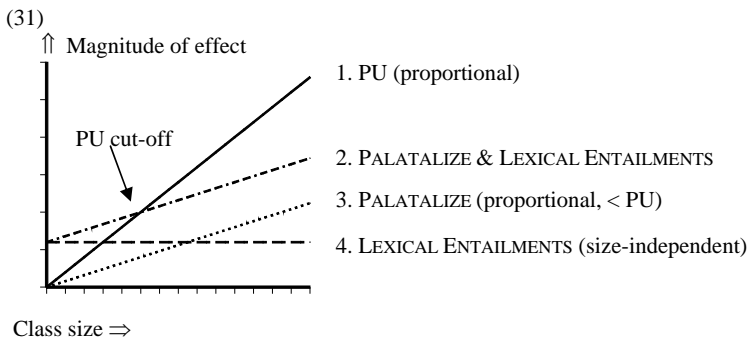
	PALATALIZE	PU	LEXICAL ENTAILMENTS
a. <i>sol-ete/solt-iamo</i> “you <sub>pl</sub> /we are used to”		*	so __-iamo ⇒ ð
a'. <i>val-ete/val-iamo</i> “you <sub>pl</sub> /we are worth”	*		va __-iamo ⇒ l
b. <i>vol-ate/vol-iamo</i> “you <sub>pl</sub> /we fly”	*		vo __-iamo ⇒ l

The reason palatalization obtains in (30a) in violation of PU is taken to be the summation of the purely phonological effect PALATALIZE (imposing assimilation to a following glide; cf. Schein & Steriade 1986) and the concurring lexical effect favoring this outcome: the entailments generated by this specific surface representation. This is parallel to the hybrid account of the pre-suffixal inserts in (10) and (11). The entailments that represent one specific lexical item can be construed as instantiating IO-FAITH in OT. The latter IO-FAITH becomes applicable to individual surface forms (here, *solt-iamo* is being faithful to itself) upon the demise of the UR, which would rather require surface allomorphs like *sol-/solt-* to have a common input (Burzio 2000, et seq.). The unifying effect of UR is reproduced here by the OO-FAITH/PU constraints, as discussed earlier. The critical contribution of the LEXICAL ENTAILMENTS in (30) is shown by the variation: Verbs that would lack such entailments are left unpalatalized (30a'). At the same time, the contribution of the phonology (PALATALIZE) is shown by the fact that palatalization is not randomly distributed: It is never found before affixes that do not instantiate the relevant environment \_\_*iV*.

Turning to class size, each of the two grammatical effects, PALATALIZE and PU, can be viewed as having magnitudes that are proportional to it. Consider that any set of stems of random shape will feature two roughly proportional subsets, one ending in *l*, like *vol-are* “fly,” the other not, like *am-are* “love.” The latter subset will then yield the entailment summation that results in PU, that is, each and every case like *am-ate/am-iamo* will generate an entailment that *X* be identical in *X-ate/X-iamo* pairs. Correspondingly, the former subset, assuming hypothetically that palatalization would apply, will similarly generate



entailments that *l/h* alternate in *...l-ate/...h-iamo*, thereby boosting the rank of PALATALIZE of (30). Hence class size maintains proportionality of the two conflicting effects. However, we know that the effect of PALATALIZE is by itself weaker than that of PU, since it fails in the absence of lexical assistance (30a'). The crucial point is that, unlike the grammatical effects, the lexical effect cannot itself match the proportionality to class size, by its very definition: it pertains to lexical singletons. No matter how large a class, there will never be more than just one verb *volare* "fly." The interaction of effects can be diagrammed as in (31).



In (31), lines 1 and 3 represent the proportionality of PU and PALATALIZE just discussed, respectively, with the latter having a smaller magnitude than the former. Line 4 represents the lexical effect, unrelated to class size. Palatalization succeeds when the summation of 4 and 3, given by line 2, exceeds the magnitude of PU. This obtains only for small classes. The intersection of lines 1 and 2 identifies the class size above which alternations will be inhibited by PU: the cut-off point.

In summary, P&B are correct that stem alternations in the minor conjugations have a substantial lexical component as in their quote in (13), since this is in fact the reason why they do not generalize to the larger *-are* conjugation. I have argued, however, that this does not disqualify such alternations from also having a substantial syntagmatic/phonological component, and that the latter is responsible for much of the distribution. In the present framework, unlike in P&B's, the lexical component is independent of the paradigmatic one (PU), and only the latter is sensitive to class size, while the former, by definition, is not. This is the reason why large classes are substantially uniform, and free of lexical effects, a prediction that cannot be reproduced by P&B's mere coindexing.

### 5. *On the difference between faithfulness and indexing*

While there is a certain affinity between the OO-FAITH relations of the present proposal and P&B's cell indexing, there is also an important difference already touched upon earlier. In OT, faithfulness constraints, like others, are violable. Indices, on the other hand, are only either the same or not. The data I discuss next further reveal the superiority of faithfulness in this regard.

Consider the parallelism between the forms of the participle and those of the preterit in (32).

(32)

Infinitive	Participle	Preterit, 1, 3, 6	Other verbs:
a. <i>chiúd-ere</i> “close”	<i>chiú-S-o</i>	<i>chiú-S-i/-e/-ero</i>	<i>accludere, attendere, chiudere, correre, corrodere, decidere, prendere, ridere, scendere, sospendere</i>
b. <i>muóv-ere</i> “move”	<i>móS-S-o</i>	<i>móS-S-i/-e/-ero</i>	<i>connettere, discutere, imprimere, muovere, riflettere</i>
c. <i>céd-ere</i> “give up”	<i>ced-út-o</i>	<i>ced-étt-i/-e/-ero</i>	<i>battere, cedere, credere, premere, ricevere, ripetere, vendere</i>

As P&B note, when the participle employs the special allomorph *-s-* as in (32a), so does the preterit in its persons 1, 3, 6. This contrasts with the pattern in (32c), where both use their respective regular allomorphs. The parallelism endures when gemination occurs as in (32b), despite the relatively idiosyncratic nature of the latter effect. P&B conclude that in (32a-b) participle and persons 1, 3, 6 of the preterit simply use the same stem: their S5. Persons 2, 4, 5 feature affixes *-ésti*, *-émmo*, *-éste*, respectively, and the stem of the infinitive (P&B's S1), for example, *chiud-ésti*—a point to which I return later.

In addition to the relationship between participle and preterit that P&B note, however, there is also one between the participle and the infinitive, which was highlighted in (26). The participial allomorph *-s-* of (32a-b) is very much like the allomorph *-t-* of (26b) *vín-t-o* and occurs only in the unstressed *-ere* (i.e., stem-stressed) conjugation. As with *-t-*, the reduced participial form *-s-* is attributable to the pressure to maintain the stem stress in the participle, the full form *-út-* causing a stress shift instead, as in (32c) (Burzio 1998, 2003). The choice between *-s-* and *-t-* is only weakly predictable from the stem form, and is—like other characteristics of the minor conjugations—partly idiosyncratic. The choice between regular allomorph *-út-* (32c) and accentual consistency (32a-b), (26b) is itself also under lexical control (Burzio 1998, 2003).

The accentual consistency of infinitive-participle pairs is in the present system a form of OO-FAITH, but, as noted before, not one that could be expressed by cell coindexing, since the two stems—*vinc-/ vin(t)-* of (26) and now also *chiud-, chius-* of (21a)—are segmentally different. Since such partial consistency cannot be expressed in P&B's system, the fact that both participial allomorphs *-t-* and *-s-* only occur in the conjugation that has stem-stressed infinitives can also not be captured. In addition to these difficulties, the *-t/-s-* variation of participles now reveals the inadequacy of the cell-indexing account of (32) even further, as shown in (33).

(33)

Infinitive	Participle	Preterit, 1, 3, 6	Other verbs
a. <i>vínc-ere</i> “win”	<i>vín-T-o</i>	<i>vín-S-i/-e/-ero</i>	<i>cogliere, estinguere, fingere, giungere, nascondere, presumere, sorgere, torcere, vincere, volgere</i>
b. <i>scrív-ere</i> “write”	<i>scríT-T-o</i>	<i>scríS-S-i/-e/-ero</i>	<i>condurre, correggere, eleggere, erigere, infliggere, scrivere</i>

What (33) shows is that the pattern of (32a-b) where gemination in the participle is reproduced in the preterit continues to hold even when the participial allomorph is *-t-*, but the preterit allomorph is *-s-*. Thus the relationship between preterit and participle is much like the one between participle and infinitive: an OO-FAITH relationship that is satisfied to the extent that other constraints permit, and not one governed by a binary choice between same or different indices. Concretely, the preterits in (33) can be accounted for by the OT ranking in (34).

(34) Partial grammar for the preterit: \**-t-* Preterit >> OO-FAITH (Participle)

The dominant constraint in (34) is simply a statement of morphological fact: The preterit does not have a *-t-* allomorph. Comparable statements will also exclude unstressed allomorphs of the inflectional affixes for persons 2, 4, 5, *ésti, -émmo, -éste*, respectively, which also break the correspondence with the participle as noted earlier. This is in contrast with the allomorphy of persons 1, 3, 6 shown in (32): stressed *-étti, -étte, -étero* alternating with unstressed *-i, -e, -ero*. The stressed allomorphs also exist in the variants *-éi, -é, -érono*, incidentally to the present discussion. When the participle uses allomorph *-t-* as in (33), the preterit will still be expected to remain faithful to it to the extent possible, namely accentually or prosodically, as indeed it does, by employing unstressed 1, 3, 6 inflections after *-s-*. Again, P&B's coindexing will not

capture such partial identity, and will therefore fail to account for the parallelism in (33) and its obvious relationship to the one in (32).

## 6. *Conclusions*

I have argued that in Italian verbal inflection lexicon, phonology, and morphology can interact in very fine-grained ways. While phonology and morphology give rise to ‘syntagmatic’ effects—surface allomorphs being conditioned by their immediate environments—such effects are under some lexical control. Yet the lexicon proves insufficient by itself, as much of the distribution is correctly captured only in syntagmatic terms.

In the extended OT framework that incorporates the representational entailments hypothesis, such interactions result from entailment summation: While either grammatical or lexical entailments may be insufficient individually, they may nonetheless be sufficient when they join forces. Entailment summation is the basis not only for such conjunctive effects, but also for constraint ranking in OT more generally. We have seen that with large classes PU effects overwhelm lexical effects, thus blocking any alternation that has a lexical basis. This is because, with larger classes, entailment summation yields larger totals.

P&B are correct in highlighting the role of paradigmatic relations, but fail to factor in the phonology responsible for many of the syntagmatic ones. In addition, their use of coindexation to express the paradigmatic relations is out of character with their violable nature, correctly expressed by OT constraints and their underlying entailments.

## REFERENCES

- Burzio, Luigi. 1998. “Multiple Correspondence”. *Lingua* 103.79-109.
- , 2000. “Cycles, Non-Derived-Environment Blocking, and Correspondence”. *Optimality Theory: Phonology, Syntax, and Acquisition* ed. by Joost Dekkers, Frank van der Leeuw, & Jeroen van de Weijer, 47-87. Oxford: Oxford University Press.
- , 2002a. “Missing Players: Phonology and the past-tense debate”. *Lingua* 112.157-199.
- , 2002b. “Surface-to-Surface Morphology: When your representations turn into constraints”. *Many Morphologies* ed. by Paul Boucher, 142-177. Somerville, MA: Cascadilla Press.
- , 2003. “Output-to-Output Faithfulness: The Italian connection”. *Lingue e Linguaggio* 1.69-104.
- , To appear. “Sources of Paradigm Uniformity”. *Paradigms in Phonological Theory* ed. by Laura J. Downing, Tracy A. Hall, & Renate Raffelsiefen. Oxford: Oxford University Press.

- DiFabio, Elvira. 1990. *The Morphology of the Verbal Infix /-isk-/ in Italian and in Romance*. Ph.D. dissertation, Harvard University.
- Halle, Morris & Alec Marantz. 1993. "Distributed Morphology". *The View From Building 20: Essays in Linguistics in honor of Sylvain Bromberger* ed. by Kenneth Hale & Samuel J. Keyser, 111-176. Cambridge, MA: MIT Press.
- Lubowicz, Anna. 1999. "Derived Environment Effects in OT". *Proceedings of the West Coast Conference on Formal Linguistics 17* ed. by Kimary N. Shahin, Susan J. Blake, & Eun-Sook Kim, 451-465. Stanford, CA: CSLI. (Available online at: <http://roa.rutgers.edu>, ROA No. 239-0198)
- Pirrelli, Vito. 2002. "Per un superamento della dicotomia lessico-grammatica. Aspetti di composizionalità 'debole' nel linguaggio". *La lessicografia bilingue tra presente e avvenire* ed. by Elena Ferrario & Virginia Pulcini, 187-204. Vercelli: Mercurio.
- Pirrelli, Vito & Marco Battista. 2000. "The Paradigmatic Dimension of Stem Allomorphy in Italian Verb Inflection". *Rivista di Linguistica* 12.307-380.
- Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar* (Report No. RuCCS-TR-2, Rutgers University Center for Cognitive Science).
- , 1997. "Optimality: From neural networks to Universal Grammar". *Science* 275.1604-1610.
- Schein, Barry & Donca Steriade. 1986. "On Geminates". *Linguistic Inquiry* 17.691-744.
- Stump, Gregory. 2001. *Inflectional Morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.
- Tantalou, Niki & Luigi Burzio. 2003. "Modern Greek Accent and Faithfulness Constraints in OT". Ms., Johns Hopkins University.

# THE ROLE OF THE L1 IN THE OVERGENERALIZATION OF CAUSATIVES IN L2 ENGLISH AND L2 SPANISH\*

MÓNICA CABRERA & MARÍA LUISA ZUBIZARRETA  
*University of Southern California*

## 0. *Introduction*

The goal of the present chapter is to investigate whether the L1 properties of lexical causatives are reflected in the interlanguage of L1 English/L2 Spanish and L1 Spanish/L2 English adult learners. We focus on the phenomenon known as overgeneralization of causatives (*\*Peter laughed the girl. "Peter caused the girl to laugh." / \*Peter arrived the girl at school late. "Peter caused the girl to arrive at school late."*).

Our findings suggest that L2 learners make use of different aspects of their L1 knowledge at different levels of proficiency. We propose that nonadvanced learners tend to focus on the L1 *constructional properties* of causatives, whereas advanced learners focus on L1 *specific lexical properties* of verb classes.

The chapter is organized as follows. In section 1, we discuss the properties of lexical causatives in English and Spanish. In section 2, previous studies on the L2 acquisition of causatives are briefly reviewed. In section 3, we present our study (hypothesis, predictions, experimental design, and group results). In section 4, the group results are discussed, and an account is proposed. Finally, in section 5, we summarize our findings and conclusions.

## 1. *Lexical causatives in Spanish and English*

### 1.1 *Some common properties of English and Spanish*

A generalization that applies to both English and Spanish is that lexical causatives are restricted to intransitives that encode change (e.g., Levin & Rappaport Hovav 1995). Verbs that typically participate in the causative alternation are change-of-state unaccusatives like *break/romper*, as exemplified in (1). There are also change-of-location unaccusatives, like Spanish *subir* "go up" and *bajar* "go down," that can alternate in transitivity, as shown in (2).

---

\* Thanks to Elaine Andersen, Stephen Crain, Toben Mintz, Silvina Montrul, William Rutherford, Mario Saltarelli, the audience of UCLA Psychobabble, and three anonymous reviewers for comments on the present study. All errors are ours.

- (1) a. Peter broke the window. / *Pedro rompió la ventana.*  
 b. The window broke. / *La ventana se rompió.*
- (2) a. *Pedro subió al niño a la mesa.*  
 “Pedro put the boy up on the table.”  
 b. *El niño se subió a la mesa.*  
 “The kid got up on the table.”

The transitive forms of such verbs are associated with the meaning given in (3).

- (3) CAUSE [Change of State/Location]

While the details of the syntactic structure associated with the meaning in (3) are not relevant to our present concerns, it is important to point out that we assume that such a structure contains a direct object of which the change of state or location is predicated (cf. Simpson 1983; Levin & Rappaport Hovav 1995; Hale & Keyser 2002; Mateu 2002). Therefore, unergative verbs like *laugh/reír* are banned from participating in the causative alternation for principled reasons, namely, they do not encode change of state or location. See (4).<sup>1</sup>

- (4) a. \*Peter laughed Mary. / \**Pedro rió a María.*  
 “Peter caused Mary to laugh.”  
 b. Mary laughed. / *María se rió.*

However, not all unaccusatives in English and Spanish that encode change of state or location appear in lexical causatives. Unaccusatives that do not alternate in transitivity include verbs of appearance (e.g., *occur/ocurrir*) and most of the inherently directed motion verbs (e.g., *arrive/llegar*); see (5) and (6). We refer to these as ‘nonalternating unaccusatives’.

- (5) a. \*Peter occurred an accident. / \**Pedro ocurrió un accidente.*  
 “Peter caused an accident to occur.”  
 b. An accident occurred. / *Un accidente ocurrió.*
- (6) a. \*Peter arrived Mary at school late. / \**Pedro llegó a María tarde a la escuela.*  
 “Peter caused Mary to arrive at school late.”  
 b. Mary arrived at school late. / *María llegó tarde a la escuela.*

<sup>1</sup> In English, but not in Spanish, certain unergatives in idiomatic usages (*bleed the patient, burp the baby, walk the dog*) can have some transitive uses. Given that these are idiomatic and highly restricted (Levin & Rappaport Hovav 1995), we assume that they do not undermine the generalization mentioned in the text.

Chierchia (1989) suggests that nonalternating unaccusatives are idiosyncratically marked for the nonlexicalization of the transitive counterpart. If Chierchia is correct, then lexical causatives with nonalternating unaccusatives, although unrealized, are not grammatically impossible in the grammar of English and Spanish.<sup>2</sup>

### 1.2A *difference between English and Spanish*

While unergatives are banned from participating in the causative alternation in both English and Spanish, the two languages differ in the following respect. As is well known, in English, but not in Spanish, manner-of-motion verbs may appear with a PP complement with a goal meaning (cf. Talmy 1985; Aske 1989); see (7) and (8).

- (7) a. The soldiers marched to the camp.  
 b. John danced to the other side of the room.
- (8) a. \**Los soldados marcharon al campamento.*  
 b. \**Juan bailó al otro lado del salón.*<sup>3</sup>
- (9) a. The soldiers marched. / *Los soldados marcharon.*  
 b. John danced. / *Juan bailó.*

Manner-of-motion verbs have different properties depending on whether they appear with a PP complement. They are unergatives when not accompanied by a PP (cf. (9)), that is, they do not necessarily encode a change of location. On the other hand, in sentences such as (7), in which the goal-denoting PP is predicated of the DP underlying object, they have unaccusative-like properties, and they encode change of location (cf. Levin & Rappaport Hovav 1995). It is therefore unsurprising that in English a manner-of-motion verb with a PP has a causative counterpart that is nonexistent in Spanish.

---

<sup>2</sup> Although nonalternating unaccusatives are generally unacceptable in lexical causatives, there are some instances of causative uses of these verbs. Carson Schütze (personal communication, 14 April 2003) provided us with the following example (uttered by an air traffic controller):

(i) We can arrive two planes an hour.

Moreover, some native speakers of Spanish are more willing to accept examples like (ii) than examples like (iii):

(ii) ?? *Si María no llega a la clase, yo la llevo.*

“If Maria does not arrive to class, I make her arrive.”

(iii) \**Si María no se ríe, yo la río.*

“If Maria does not laugh, I make her laugh.”

<sup>3</sup> Sentence (8b) is acceptable in Spanish only with a locative interpretation. With a directional (change-of-location) interpretation, it is unacceptable.



- (10) a. The general marched the soldiers to the camp.  
 b. John danced Mary across the room.<sup>4</sup>
- (11) a. \**El general marchó a los soldados al campamento.*  
 b. \**Juan bailó a María al otro lado del salón.*

Verbs of manner of motion require a PP complement that encodes change of location in order to be acceptable in lexical causatives. In effect, although the English causatives in (12) are perhaps not as strongly unacceptable as their Spanish counterparts in (13), there is a clear contrast between (10) and (12).

- (12) a. ??The general marched the soldiers.  
 b. \* John danced Mary.
- (13) a. \* *El general marchó a los soldados.*  
 b. \* *Juan bailó a María.*

### 1.3 A summary

We summarize in (14), (15), and (16) the similarities and differences between English and Spanish lexical causatives. These can be described along two dimensions: the general properties of the construction and specific properties of particular subclasses of verbs. We refer to the former as ‘constructional properties’ and to the latter as ‘specific lexical properties.’

(14) General property of the lexical causative construction:

In English and Spanish, the structure [DP V DP (PP)] can be associated with the meaning in (3) iff the verb and/or the complement PP encodes a change of state or location that is predicated of the object. (Unaccusatives that encode change therefore satisfy the general constructional requirement. Unergatives do not encode change and therefore they do not satisfy the general constructional specification.)

Specific properties of subclasses of verbs:

- (15) In English and Spanish, a subset of change-of-state/-location unaccusatives (i.e., verbs of appearance and most of the inherently directed motion verbs) fails to enter the causative alternation.
- (16) In English, a manner-of-motion verb can appear in the causative construction in the context of a PP that encodes change of location. In Spanish, a manner-of-motion verb can never appear in the causative construction.

In our study, we investigate the role of L1 transfer in the L2 acquisition of lexical causatives. More specifically, we study how the aforementioned properties of lexical causatives in English and Spanish are reflected in the interlanguage of L2 learners. Before we turn to our study, we briefly review

---

<sup>4</sup> An anonymous reviewer points out that, although sentences in (10) are acceptable in English, they are highly marked.

some previous studies on the L2 acquisition of lexical causatives in Spanish and English.

## 2. *Previous studies on L2 causatives*

Overgeneralization of causatives (verbs used with a causative meaning that is unacceptable in the target language) has been studied in adult L2 acquisition of English and Spanish by Moore (1993), Montrul (1997, 1999, 2001a, 2001b), and Cabrera and Zubizarreta (2003a). Montrul (1997, 1999, 2001b) found that intermediate L2 learners of English (L1 Spanish) and Spanish (L1 English) accepted overgeneralized causatives with nonalternating unaccusatives (cf. (5a) and (6a)) and unergatives (cf. (4a)) without making a significant distinction between these verb classes. The author concluded that the phenomenon is independent of L1 transfer, and that it is due to the learner's lack of knowledge of the relevant lexicosemantic features that determine which verbs can alternate in transitivity. More precisely, Montrul proposed that, due to their incomplete lexical L2 knowledge at early stages of acquisition, learners resort to a default transitive lexicosemantic template (*NP CAUSE NP BECOME verb*) that they use irrespective of verb type.

Cabrera and Zubizarreta (2003a) partially replicated Montrul's (1997, 1999, 2001b) study with L1 English/L2 Spanish learners at different levels of proficiency in order to further investigate whether L2 learners are sensitive to the nonalternating unaccusative/unergative distinction at some point of the acquisition process. Differently from Montrul (1997, 1999, 2001b), it was found that beginner and intermediate learners overgeneralized causatives significantly more with nonalternating unaccusatives than with unergatives.<sup>5</sup> This result is not surprising given that in both the L1 and the L2 the intransitives that participate in the causative alternation are unaccusatives. The preference for unaccusatives in causative structures was therefore explained as

---

<sup>5</sup> The group results do not imply that all learners overgeneralized causatives more with unaccusatives than with unergatives. The individual analysis in Cabrera and Zubizarreta (2003a) indicated that, among the learners who overgeneralized causatives, the largest group was formed by those that did so only or mostly with unaccusatives (35 subjects). However, there was a small group (2 subjects) that overgeneralized causatives with unaccusatives and unergatives equally. Two strategies were identified, summarized in (i). Learners who used a syntactic strategy (i-a) overgeneralized with both unergatives and unaccusatives, and those who used a lexicosyntactic strategy (i-b) overgeneralized only with unaccusatives.

(i) a. Syntactic strategy: The lexical information of particular verb classes is ignored.

The following general form-meaning correspondence generalization is used:

Surface template  $NP_1 V NP_2 \leftrightarrow [NP_1 CAUSE [NP_2 BECOME PRED]]$

b. Lexicosyntactic strategy: The correspondence between the previous surface syntactic template and meaning is restricted to a lexicosyntactic class, namely, unaccusatives.

a case of overgeneralization of the lexicosyntactic properties of alternating unaccusatives (*break/romper*) to nonalternating unaccusatives (*arrive/llegar*). Contrary to Montrul's (1997, 1999, 2001b) proposal, this finding is compatible with a transfer analysis in which learners exploit selected L1 properties in constructing their interlanguage.

To our knowledge, the only study to address the L2 acquisition of lexical causatives with manner-of-motion verbs (cf. (10) and (11)), is Montrul (2001a). The participants in this study were L1 English speakers learning Spanish and L1 Spanish speakers learning English, at the intermediate level of proficiency. The results showed L1 transfer effects. L1 English speakers seemed to be constrained by their L1, in that they tended to accept sentences like (11) in L2 Spanish. On the other hand, L1 Spanish speakers were also constrained by their L1, since they tended to reject sentences like (10) in L2 English. In other words, L1 English speakers overgeneralized causatives with verbs of manner of motion, whereas L1 Spanish speakers undergeneralized causatives with this class of verbs. Montrul concluded that L1 transfer applies in certain cases but not in others. L1 effects would be more pervasive with language-specific alternations (such as causativization with [manner-of-motion verbs + PP] in English) than with alternations of a more universal scope (like the causative alternation with unaccusatives of change of state/location). These apparently contradictory results point to the need for a more comprehensive study on the overgeneralization of causatives across different levels of proficiency, which simultaneously tests for lexical causative structures in which English and Spanish are similar and those in which they are different (cf. (14)-(16)). A study of this nature should elucidate what role the L1 plays in the L2 acquisition of lexical causatives, if any, and at what level of proficiency. This is precisely the objective of the present study.

### 3. *The present study*

Our study is a partial replication/extension of Montrul (1997, 1999, 2001a, 2001b) with two experimental groups, L1 English/L2 Spanish and L1 Spanish/L2 English speakers, across different levels of proficiency. In order to investigate the role of the L1 in the L2 acquisition of lexical causatives, we tested structures in which English and Spanish have the same properties (cf. (14)-(15)), and those in which these languages behave differently (cf. (16)).

#### 3.1 *Hypothesis and predictions*

The central hypothesis of the present study is that the properties of the L1 determine which verb classes appear in lexical causatives in the interlanguage.

If there is transfer of the constructional properties (cf. (14)), the following prediction can be made:

- (17) Both experimental groups should accept lexical causatives significantly more with nonalternating unaccusatives (e.g., *arrive/llegar*) and [manner-of-motion verbs + PP] (e.g., [*dance/bailar* + PP]) than with unergatives (e.g., *laugh/reír*) and manner-of-motion verbs without a PP (e.g., *dance/bailar*).

If there is transfer of specific lexical properties (cf. (15)), the following predictions are made:

- (18) a. Both experimental groups should reject lexical causatives with nonalternating unaccusatives (e.g., *arrive/llegar*).  
 b. The L1 English/L2 Spanish group should accept lexical causatives with [manner-of-motion verbs + PP] (e.g., [*bailar* + PP]).  
 c. The L1 Spanish/L2 English group should reject lexical causatives with [manner-of-motion verbs + PP] (e.g., [*dance* + PP]).

Note that the predictions in (17) and those in (18) are to a great extent mutually exclusive. In effect, this is the case because the effects of the specific lexical constraints override the effects of the constructional constraints. Therefore, we can tease apart transfer of the two types of properties if and only if they are at play at different levels of proficiency.

### 3.2 *Experimental design*

3.2.1 *Participants*. A total of 153 adults participated in the study. There were two experimental and two control groups. The L1 English/L2 Spanish experimental group consisted of 43 students in the Spanish Basic Language Program at the University of Southern California (mean age = 19.26), tested in Los Angeles, California. The L1 Spanish/L2 English experimental group consisted of 73 students in the English Language Program at the Pontificia Universidad Católica del Perú (mean age = 21.52), tested in Lima, Perú. The control groups consisted of 18 native-Spanish-speaking students at the Pontificia Universidad Católica del Perú (mean age = 24.94), tested in Lima, Perú, and of 19 native-English-speaking students at the University of Southern California (mean age = 18.95), tested in Los Angeles, California.

The L2 proficiency level of the experimental groups was measured using a cloze test, which included 3 paragraphs with a total of 75 blank spaces (1 blank space every 5 words), and which was corrected with the acceptable word criterion (Butler 1980; Schiarone & Schoorl 1989). For each experimental group, 3 levels of proficiency were obtained. In the L1 English/L2 Spanish

experimental group, there were 17 beginners, 16 intermediates, and 10 advanced learners. In the L1 Spanish/L2 English experimental group, we found 22 beginners, 28 intermediates, and 23 advanced learners. The proficiency levels were significantly different from each other and from their corresponding control group ( $p < .0001$ ) in terms of their mean score on the cloze test.

3.2.2 *Tests.* Besides the cloze proficiency test, a verb-translation task (VTT) and an acceptability judgment test (AJT) were used, testing a total of 26 verbs. In the VTT, subjects were asked to translate a list of verbs in the L2 (cf. Table 1) into their L1. The purpose of this test was to determine whether the subjects knew the idiosyncratic meanings of verbs; only the AJT responses corresponding to correctly translated verbs were used in computing results. The verb classes tested in this task were nonalternating unaccusatives, unergatives, manner-of-motion verbs, and alternating unaccusatives.

The purpose of the AJT, based on Montrul's (1997)'s design, was to see whether the subjects accepted the verb classes tested in the VTT in the lexical causative configuration. A total of 32 test items and 15 filler sentences were used. The verbs of manner of motion were tested first without a PP, and later with a directional PP.<sup>6</sup> Each sentence was accompanied by a picture to ensure the correct interpretation.<sup>7</sup> Subjects were instructed to rate the sentences focusing on grammaticality/acceptability in the target language, according to a 7-point Likert scale, from -3 (completely unacceptable) to +3 (completely acceptable). Subjects were also instructed to use the value 0 in case they thought the sentence was neither good nor bad, and to leave the question blank if they were not sure about the acceptability of the sentence. Examples of tested sentences are shown in Table 2.

---

<sup>6</sup> The test included four parts. In part 1, verbs appeared in an intransitive configuration. In parts 2 and 3, verbs were tested in a lexical causative configuration (manner-of-motion verbs without a PP were tested in part 2, and with a directional PP in part 3). Finally, part 4 included all verbs in periphrastic causatives. In this chapter, we report the results on lexical causatives only (parts 2 and 3).

<sup>7</sup> Thanks to Silvina Montrul for allowing us to use her testing materials in our pilot study. For the picture design of this study, we thank Pierre Canueil and Mabel Amaya de Beas.

Nonalternating Unaccusatives	Unergatives	Manner-of-Motion Verbs	Alternating Unaccusatives
caer / fall	fumar / smoke	desfilarse / parade	abrir / open
ir / go	ladrar / bark	marchar / march	cerrar / close
llegar / arrive	llorar / cry	correr / run	cocer / cook
entrar / enter	luchar / fight	saltar / jump	quemar / burn
aparecer / appear	acampar / camp	bailar / dance	romper / break
ocurrir / happen	reír / laugh	volar / fly	secar / dry
salir / leave			
venir / come			

Table 1: *Tested verbs in the VTT and in the AJT*

Nonalternating Unaccusatives	*El padre <u>llegó</u> a la niña tarde a la escuela. *The father <u>arrived</u> the girl at school late.
Unergatives	*Pedro <u>rió</u> a Juan. *Peter <u>laughed</u> John.
Manner-of-Motion Verbs	*El general <u>marchó</u> a los soldados. ??The general <u>marched</u> the soldiers.
Manner-of-Motion Verbs + PP	*El general <u>marchó</u> a los soldados <u>al campamento</u> . The general <u>marched</u> the soldiers <u>to the camp</u> .
Alternating Unaccusatives	Pedro <u>rompió</u> la ventana. Peter <u>broke</u> the window.

Table 2: *Examples of lexical causatives tested in the AJT*

### 3.3 Group results

In this section, we report group results corresponding to verbs only in the lexical causative configuration for both of our experimental groups: L1 English/L2 Spanish and L1 Spanish/L2 English. For lack of space, we do not present the results of the individual analysis (see Cabrera & Zubizarreta 2003b).

3.3.1 *Alternating unaccusatives*. Figure 1 shows the acceptability means for alternating unaccusative verbs (*break/romper*) in lexical causatives.<sup>8</sup> Both experimental groups correctly accepted these sentences. A one-way ANOVA that compared each experimental group to their corresponding control group indicated that there was no significant difference between them, that is, L2 learners behaved like the controls with respect to alternating unaccusatives.

<sup>8</sup> Error bars in Figures 1 through 5 represent the standard error of the mean.

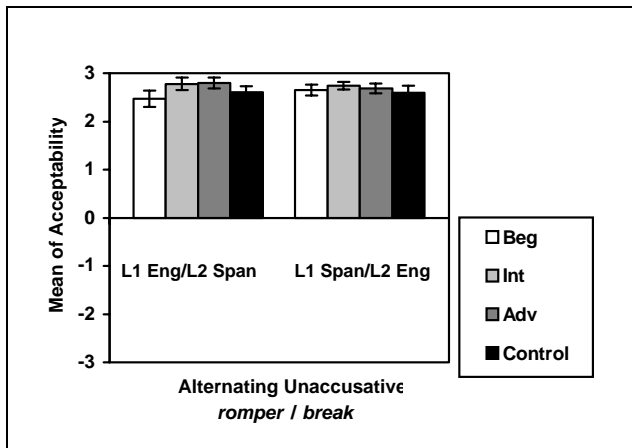


Fig. 1: Lexical causatives with alternating unaccusatives

On the other hand, in order to determine whether the experimental groups were sensitive to the difference between lexical causatives with alternating unaccusatives, which are acceptable in both L2s (Spanish and English), and lexical causatives with nonalternating unaccusatives, unergatives, and manner-of-motion verbs (with or without PP), we ran paired-sample *t* tests comparing the means of acceptability for these lexical causatives across proficiency levels. L1 English/L2 Spanish learners of all proficiency levels significantly preferred lexical causatives with alternating unaccusatives (Figure 1) to those with the other verb classes (Figures 2 and 4;  $p < .001$  for all comparisons). The L1 Spanish/L2 English group showed the same preference, that is, alternating unaccusatives in lexical causatives (Figure 1) were significantly preferred to the other verb classes (Figures 3 and 5;  $p < .0001$  for all comparisons).

3.3.2 *Comparing nonalternating unaccusatives, manner-of-motion verbs with PP, and unergatives.* Figure 2 illustrates the acceptability means of the lexical causative form of nonalternating unaccusatives,<sup>9</sup> manner-of-motion verbs with PP, and unergatives for the L1 English/L2 Spanish group. Using one-way ANOVAs, it was found that learners significantly accepted more lexical causatives with unaccusatives and manner-of-motion verbs with PP than the

<sup>9</sup> The means of acceptability for nonalternating unaccusatives, unergatives, and manner-of-motion verbs with PP (Figures 2 and 3) have a value close to 0. These means reflect that some learners accepted lexical causatives with the aforementioned verb classes and others did not. See Cabrera and Zubizarreta (2003b) for the individual analysis of the data.

control group ( $p < .0001$ ). As for unergatives, beginners and intermediates tended to reject them, but did so significantly less often than the advanced and the control groups ( $p < .0001$ ). The L1 English/L2 Spanish group showed a tendency to accept lexical causatives that are unacceptable in the L2.

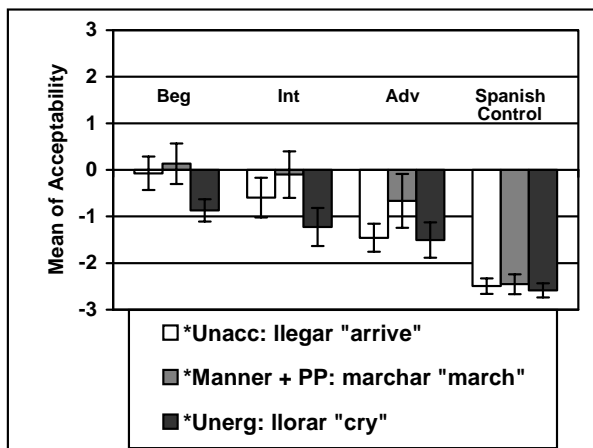


Fig. 2: L1 English/L2 Spanish: Lexical causatives with nonalternating unaccusatives, manner-of-motion verbs with PP, and unergatives

In order to compare these verb classes over proficiency levels, paired-sample  $t$  tests were performed. For beginners and intermediates, there was a significant difference between the means for unaccusatives and manner-of-motion verbs with PP, and the mean for unergatives ( $p < .05$ ). Beginners and intermediates preferred lexical causatives with verbs that encode change of state/location (unaccusatives and manner-of-motion verbs with PP) to the ones with verbs that do not (unergatives). No significant difference between the means of unaccusatives and manner-of-motion verbs with PP was found. These learners treated verbs that encode change of state/location similarly. For the advanced group, there was no significant difference between unaccusatives and unergatives. Advanced learners equally rejected these verb classes. However, there was a significant difference between the means for unaccusatives and unergatives, on the one hand, and the mean for manner-of-motion verbs with PP, on the other ( $p < .05$ ). The advanced group equally rejected nonalternating unaccusatives and unergatives, and significantly preferred lexical causatives with manner-of-motion verbs with PP, which are acceptable in their L1.



Figure 3 shows the acceptability means for nonalternating unaccusatives, manner-of-motion verbs with PP, and unergatives for the L1 Spanish/L2 English group. Using one-way ANOVAs, we found that learners behaved significantly differently from the control group for unaccusatives and unergatives: they accepted them significantly more than the control group ( $p < .0001$ ), although, as we will see, there was a significant difference between unaccusatives and unergatives. As for manner-of-motion verbs with PP, beginners did not behave significantly differently from the control group, whereas intermediates and advanced did ( $p < .0001$ ). We can also say in this case that the L1 Spanish/L2 English group showed a tendency to accept lexical causatives that are unacceptable in the L2, with the exception of manner-of-motion verbs with PP, in the case of beginners.

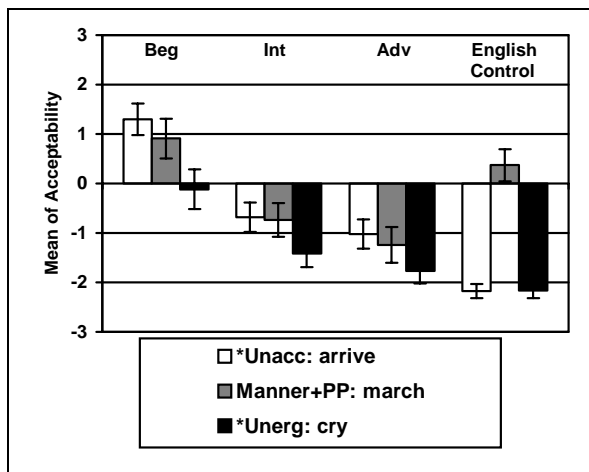


Fig. 3: L1 Spanish/L2 English: Lexical causatives with nonalternating unaccusatives, manner-of-motion verbs with PP, and unergatives

Paired-sample  $t$  tests by proficiency level were performed to compare these verb classes. For beginners ( $p < .05$ ) and intermediates ( $p < .005$ ), there was a significant difference between the means for unaccusatives and manner-of-motion verbs with PP, and the mean for unergatives. Beginners and intermediates preferred lexical causatives with verbs encoding change of state/location (unaccusatives and manner-of-motion verbs with PP) to the ones with verbs that do not (unergatives). No significant difference was found between unaccusatives and manner-of-motion verbs with PP. These learners treated verbs that encode change of state/location similarly. For the advanced

learners, there was a significant difference between unaccusatives and manner-of-motion verbs with PP, on the one hand, and unergatives, on the other ( $p < .05$ ). The advanced group rejected to a lesser extent lexical causatives with verbs encoding change than those with unergative verbs. There was no significant difference between unaccusatives and manner-of-motion verbs with PP. In contrast to the L1 English/L2 Spanish advanced group, they equally rejected lexical causatives with unaccusatives and manner-of-motion verbs with PP, both of which are unacceptable in their L1.

3.3.3 *Comparing manner-of-motion verbs with and without PP.* The acceptability means for manner-of-motion verbs with and without PP were compared in order to further explore whether L2 learners were sensitive to the fact that, when the goal-denoting PP is present, these verbs encode change of location. The means for the L1 English/L2 Spanish group are presented in Figure 4. Paired-sample  $t$  tests compared the means across levels of proficiency. A significant difference was found between the means of manner-of-motion verbs with PP and without PP for the beginner ( $p < .05$ ) and the advanced ( $p < .005$ ) groups. For intermediates, there was no significant difference, but a trend ( $p = .09$ ) was found. Namely, the L1 English/L2 Spanish group preferred lexical causatives with verbs of manner when the PP was present. In other words, these learners preferred verbs of manner of motion when they encoded a change of location. This result also shows that the learners more readily accepted lexical causatives that are allowed in their L1.

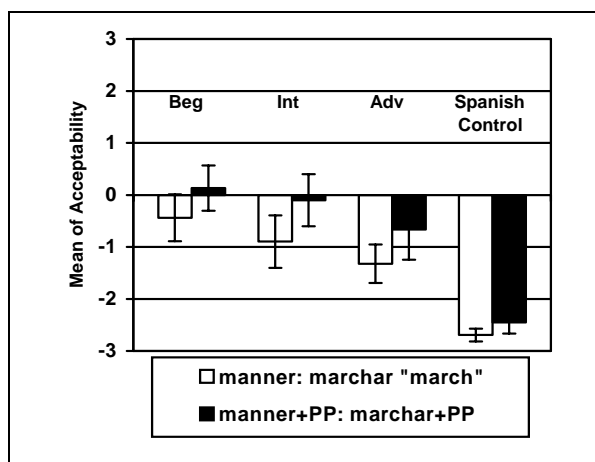


Fig. 4: L1 English/L2 Spanish: Lexical causatives with manner-of-motion verbs with and without PP

The means for verbs of manner of motion with and without a directional PP for the L1 Spanish/L2 English group are illustrated in Figure 5. We used paired-sample *t* tests to perform the same comparisons as for the L1 English/L2 Spanish group. We found a significant difference between the means for beginners ( $p < .05$ ) and intermediates ( $p < .005$ ). There was no significant difference for the advanced group. Beginners preferred and intermediates rejected to a lesser extent manner-of-motion verbs with a directional PP in lexical causatives, whereas advanced learners equally rejected these verbs with or without a PP.

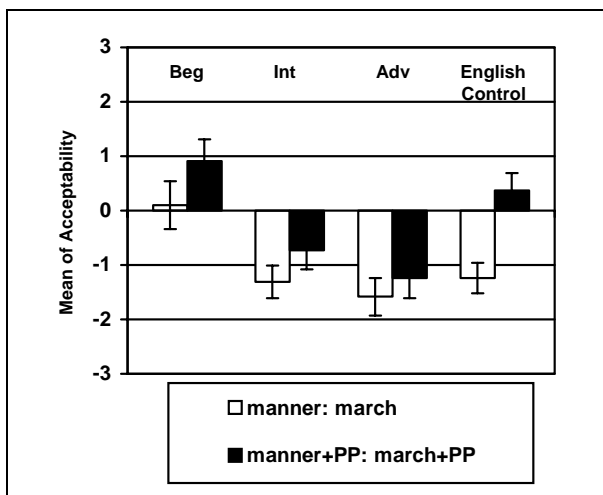


Fig. 5: L1 Spanish/L2 English: Lexical causatives with manner-of-motion verbs with and without PP

### 3.4 Generalizations

The results presented in section 3.3 lead us to the following generalizations with respect to how the experimental groups treated lexical causatives with nonalternating unaccusatives, unergatives, and manner-of-motion verbs (+ PP):

(19) a. Generalization 1:

Both experimental groups behaved in a similar way at the beginner and intermediate levels of proficiency. They preferred (or rejected to a lesser extent) verbs encoding change of state or location (nonalternating unaccusatives and manner-of-motion verbs with PP) to unergatives. The different classes of verbs of change of state/location were treated equally. They also accepted manner-of-motion verbs more often when a PP was present than when it was not.

b. Generalization 2:

The experimental groups behaved differently at the advanced level of proficiency. The L1 English/L2 Spanish group rated manner-of-motion verbs with PP significantly higher than nonalternating unaccusatives. The L1 Spanish/L2 English group did not show a preference for either of these verb classes: they rejected both equally.

On the other hand, the L1 English/L2 Spanish group rejected nonalternating unaccusatives and unergatives equally, whereas the L1 Spanish/L2 English group still rejected nonalternating unaccusatives less often than unergatives.

In the next section, we present our interpretation and analysis of the data, and discuss whether the predictions (cf. (17) and (18)) were borne out.

#### 4. *Analysis and discussion*

The generalizations in (19) confirmed our hypothesis and its corresponding predictions. Our central hypothesis was that L1 properties determine which verb classes appear in lexical causatives in the interlanguage. If there is transfer of the constructional properties (cf. (14)), both experimental groups were expected to accept nonalternating unaccusatives (e.g., *arrive/llegar*) and manner-of-motion verbs with PP (e.g., *dance/bailar* + PP) in lexical causatives more than unergatives (e.g., *laugh/reír*) and manner-of-motion verbs without PP (e.g., *dance/bailar*). This prediction was borne out for the beginner and intermediate levels of proficiency of both experimental groups (cf. (19a)), but not for the advanced group (cf. (19b)).

On the other hand, if there is transfer of specific lexical properties (cf. (15)-(16)), both experimental groups were expected to reject lexical causatives with nonalternating unaccusatives (cf. (18a)). However, the experimental groups were predicted to behave differently in that the L1 English/L2 Spanish group was expected to accept manner-of-motion verbs with PP (*bailar* + PP) in lexical causatives (cf. (18b)), whereas the L1 Spanish/L2 English group was expected to reject those verbs in that configuration (cf. (18c)). These predictions held for the advanced level of proficiency (cf. (19b)) in both experimental groups, but not for beginners or intermediates (cf. (19a)). It was found that, when recovering from the overgeneralization of causatives with nonalternating unaccusatives, the L1 English advanced group still showed some preference for overgeneralized causatives with manner-of-motion verbs with PP, while the L1 Spanish advanced group undergeneralized causatives in that context. We should point out that the advanced learners in both experimental groups rejected not only nonalternating unaccusatives in lexical causatives, but also unergatives, which suggests that they transfer constructional properties at this level of proficiency as well (cf. (14)).

We put forth the proposal that overgeneralization (and undergeneralization) of causatives can be reduced to transfer of different L1 properties (i.e. constructional and specific lexical) at different levels of proficiency. Our central claim is that learners gradually make use of different aspects of their L1 knowledge in order to analyze the L2 input. More specifically, our data may be interpreted as suggesting that learners focus first on constructional properties of lexical causatives, and later they also make use of specific lexical properties of verb classes.

#### 4.1 *Stage 1: Focus on the construction*

Our first proposal is that, at earlier stages of acquisition, the L2 learner focuses on the properties of the construction (cf. (14)). Namely, a causative meaning can be associated with a transitive construction if and only if it is licensed by a lexical item (verb or preposition) that encodes a change of state or location. English and Spanish are alike with respect to the general licensing of the causative construction. Therefore, L1 English/L2 Spanish learners and L1 Spanish/ L2 English learners behave similarly at the earlier stage of acquisition. More precisely:

- (20) a. Both groups of L2 learners accept more lexical causatives with verbs encoding change, that is, nonalternating unaccusatives (*arrive/llegar*), and manner-of-motion verbs with PP (*dance/bailar* + PP).
- b. Both groups accept to a lesser extent lexical causatives with verbs that do not encode change, that is, unergatives (*laugh/reír*) and manner-of-motion verbs without a directional PP.

#### 4.2 *Stage 2: Focus on specific lexical properties of the verb as well*

We propose that, at a later stage of acquisition, the L2 learner restricts the application of the constructional properties to the verb classes that are allowed in lexical causatives in the L1 (cf. (15)-(16)). In other words, specific lexical properties are put to use together with the constructional properties. Therefore, L1 English learners of L2 Spanish and L1 Spanish learners of L2 English will behave similarly in some respects and differently in others. More precisely:

- (21) a. Both groups reject lexical causatives with nonalternating unaccusatives (*arrive/llegar*) and unergatives (*laugh/reír*).
- b. L1 English/L2 Spanish learners incorrectly accept lexical causatives with manner-of-motion verbs with PP (*dance* + PP). L1 Spanish/L2 English learners incorrectly reject lexical causatives in that context.

Why do L2 learners transfer different L1 properties at different levels of proficiency? Following VanPatten (1996), our view is that L2 learners

gradually use different properties of their L1 in order to derive intake from L2 input. 'Intake,' a term coined by Corder (1981), is defined by VanPatten (1996:10) as "the subset of filtered input that serves as the data for accommodation by the developing system." In other words, L2 learners do not process L2 input fully, but rather filter it in order to derive intake, which is then used to build the interlanguage. VanPatten also claims that intake formation is constrained by attention, which limits the L2 learner's capacity to deal with stimuli. Our data suggest that L2 learners are able to use more aspects of their L1 knowledge to analyze the L2 input in a more detailed manner only when they are more proficient, when their interlanguage is thus more developed.

On the other hand, why are constructional properties transferred before specific lexical properties? We suggest that the preference for the former over the latter is also related to input-processing mechanisms. According to syntax first (or two-stage) models of sentence comprehension or parsing (Frazier & Rayner 1982; Ferreira & Clifton 1986; Ferreira & Henderson 1990, among others), speakers tend to process sentences making use of syntactic frames. In our study, L2 learners paid attention first to lexical items that can license the causative construction (verb or preposition encoding change), and preferred lexical causatives including a licenser of the construction. In other words, at earlier stages of acquisition, L2 learners seem to focus on the syntactic properties of some lexical items that are necessary to license the causative construction.

This brings up the question of why advanced learners seem to still be influenced by their L1, taking into consideration that they have been exposed to more L2 input. Our data shows that some advanced learners still overgeneralized (or undergeneralized) causatives with manner-of-motion verbs with a directional PP. In the case of L1 English/L2 Spanish learners, the recovery from overgeneralization is presumably more difficult because some form of negative evidence (direct or indirect) would be necessary. The L1 Spanish/L2 English advanced learners' situation would be more easily overcome in that they would only require positive evidence to recover from undergeneralization. A possibility is that our subjects at the advanced proficiency level are not advanced enough to have completely overcome overgeneralization (and undergeneralization) errors. This issue deserves further investigation involving learners of an even higher proficiency level.

## 5. *Conclusion*

This chapter argues that the phenomenon of overgeneralization of causatives in L2 acquisition can be reduced to L1 transfer of different aspects

of L1 knowledge. At earlier stages of L2 acquisition (beginner and intermediate proficiency), the constructional meaning of lexical causatives seems to trigger the overgeneralization of causatives, in particular, with verbs encoding change of state or location. However, at the advanced proficiency stage, when recovering from overgeneralization, L1 specific lexical constraints come into play. The results from our study suggest that L2 learners transfer L1 properties gradually, and not all at once.

An argument that might be put forth against the transfer analysis of overgeneralized causatives is that the phenomenon also exists in L1 acquisition (Lord 1979; Bowerman 1982; Braine et al. 1990; Pinker 1989). Nevertheless, the research to date does not show that there is more overgeneralization with unaccusatives than with unergatives in child language (Bowerman 1996). We therefore tend to think that overgeneralization of causatives in adult L2 and child L1 acquisition does not necessarily have the same source. It would seem that children who overgeneralize causatives freely associate a transitive structure with a causative meaning (irrespective of verb type). This would be akin to the purely syntactic strategy that some of the L2 learners resort to and that the individual analysis in Cabrera and Zubizarreta (2003a) revealed (see fn. 4).

Finally, this chapter demonstrates that research in L2 acquisition can shed light on the L1 grammar as well as on the interlanguage grammar. In the case at hand, the interlanguage data shows that ‘change of state or location’ is a relevant semantic factor in the characterization of the verb classes that underlie the lexical causative construction in L1 English and L1 Spanish.

## REFERENCES

- Aske, Jon. 1989. “Motion Predicates in English and Spanish: A closer look”. *Proceedings of the Fifteenth Annual Meeting of the Berkeley Linguistics Society (BLS 15)* ed. by Kira Hall, Michael Meacham, & Richard Shapiro, 1-14. Berkeley, CA: Berkeley Linguistics Society.
- Bowerman, Melissa. 1982. “Evaluating Competing Linguistic Models with Language Acquisition Data: Implications of developmental errors with causative verbs”. *Quaderni di Semantica* 3.5-66.
- , 1996. “Argument Structure and Learnability: Is a solution in sight?”. *Proceedings of the Twenty-Second Annual Meeting of the Berkeley Linguistics Society (BLS 22)* ed. by Janice M. Johnson, Matthew L. Juge, & Jeri L. Moxley, 454-468. Berkeley, CA: Berkeley Linguistics Society.
- Braine, Martin D, S., Ruth E. Brody, Shalom M. Fisch, Mara J. Weisberger, & Monica Blum. 1990. “Can Children Use a Verb without Exposure to its Argument Structure?”. *Journal of Child language* 17.313-342.

- Butler, Frances. 1980. "Cloze as an Alternative Method of ESL Placement and Proficiency Testing". *Research in Language Testing* ed. by John Oller & Kyle Perkins, 121-128, Boston: Newbury House.
- Cabrera, Mónica & María Luisa Zubizarreta. 2003a. "On the Acquisition of Spanish Causative Structures by L1 Speakers of English", *Proceedings of the 2002 Generative Approaches to Second Language Acquisition Conference (GASLA-2002): L2 links* ed. by Juana M. Licerias, Helmut Zobl, & Helen Goodluck, 24-33. Somerville, MA: Cascadilla Proceedings Project.
- , 2003b. "Overgeneralized Causatives in L2 English and L2 Spanish". Ms., University of Southern California.
- Chierchia, Gennaro. 1989. "A Semantics for Unaccusatives and Its Syntactic Consequences". Ms., Cornell University.
- Corder, S. Pit. 1981. *Error Analysis and Interlanguage*. Oxford: Oxford University Press.
- Ferreira, Fernanda & Charles Clifton, Jr. 1986. "The Independence of Syntactic Processing". *Journal of Memory and Language* 25.348-368.
- Ferreira, Fernanda & John M. Henderson. 1990. "Use of Verb Information in Syntactic Parsing: Evidence from eye movements and word-by-word self-paced reading". *Journal of Experimental Psychology: Learning, Memory, and Cognition* 16.555-568.
- Frazier, Lyn & Keith Rayner. 1982. "Making and Correcting Errors During Sentence Comprehension: Eye movements in the analysis of structurally ambiguous sentences". *Cognitive Psychology* 14.178-210.
- Hale, Ken & Samuel J. Keyser. 2002. *Prolegomenon to a Theory of Argument Structure*. (Linguistic Inquiry Monograph). Cambridge, MA: MIT Press.
- Levin, Beth & Malka Rappaport Hovav. 1995. *Unaccusativity at the Syntax Semantics Interface* (Linguistic Inquiry Monograph). Cambridge, MA: MIT Press.
- Lord, Carol. 1979. "Don't you fall me down: Children's generalizations regarding cause and transitivity". *Papers and Reports on Child Language Development (PRCLD 17)* ed. by Eve V. Clark, 81-19. Stanford, CA: Stanford University Department of Linguistics.
- Mateu, Jaume. 2002. "Arguing Our Way to the Direct Object Restriction on English Resultatives". Ms., Universitat Autònoma de Barcelona.
- Montrul, Silvina. 1997. *Transitivity Alternations in Second Language Acquisition: A cross-linguistic study of English, Spanish, and Turkish*. Ph.D. dissertation, McGill University.
- , 1999. "Causative Errors with Unaccusative Verbs in L2 Spanish". *Second Language Research* 15:2.191-219.
- , 2001a. "Agentive Verbs of Manner of Motion in Spanish and English as Second Languages". *Studies in Second Language Acquisition* 23.145-151.
- , 2001b. "Causatives and Transitivity in L2 English". *Language Learning* 51:1.51-106.
- Moore, Miriam. 1993. *Second Language Acquisition of Lexically Constrained Transitivity Alternations: Acquisition of the causative alternation by second language learners of English*. Ph.D. dissertation, University of South Carolina.
- Pinker, Steven. 1989. *Learnability and Cognition: The acquisition of argument structure*. Cambridge, MA: MIT Press.
- Schiarone, A. & J. School. 1989. "The Cloze Test: Or why small isn't always beautiful". *Language Learning* 39:3.415-438.
- Simpson, Jane. 1983. "Resultatives". *Papers in Lexical-Functional Grammar* ed. by Lori S. Levin, Malka Rappaport, & Annie Zaenen, 143-157. Bloomington: Indiana University Linguistics Club.



- Talmy, Leonard. 1985. "Lexicalization Patterns: Semantic structures in lexical forms".  
*Language Typology and Syntactic Description III: Grammatical categories and the lexicon*, ed. by Timothy Shopen, 57-149. Cambridge: Cambridge University Press.
- VanPatten, Bill. 1996. *Input Processing and Grammar Instruction: Theory and research*.  
Norwood, NJ: Ablex.

# THE INCHOATIVE INTERPRETATION OF THE *IMPERFECTO*\*

ALICIA CIPRIA  
*University of Alabama*

## 0. *Introduction*

The inchoative or inceptive reading of the *imperfecto*, when the beginning of an event is signaled to start taking place after the past utterance time, as in the Spanish example (1), is one of the interpretations of the *imperfecto* that has not received much analysis in either the traditional or the formal literature:

- (1) *Frida ensayaba el libreto en una hora.*  
Frida rehearse<sub>IMPF</sub> the libretto in an hour  
“Frida rehearsed/used to rehearse/was rehearsing/intended/was to rehearse the libretto in an hour.”  
(telic: inchoative; or atelic: habitual, progressive, or futurate)

Even though (1) may be typically found in contexts such as (2a,b), it does bring about interesting issues regarding the interaction with adverbials, the nature of truly inchoative readings and coercion phenomena.<sup>1</sup>

- (2) a. *Frida dijo que ensayaba el libreto en una hora.*  
Frida say<sub>PRET</sub> that *pro*-rehearse<sub>IMPF</sub> the libretto in an hour  
“Frida said that she rehearsed (was rehearsing) the libretto in an hour.”  
b. *Frida comenzó a sudar nerviosamente. Ensayaba el libreto en una hora.*  
Frida start<sub>PRET</sub> to sweat nervously. *pro*-rehearse<sub>IMPF</sub> the libretto in an hour.  
“Frida started to sweat nervously. She rehearsed (was rehearsing) the libretto in an hour.”<sup>2</sup>

---

\* I thank three anonymous reviewers for their excellent comments; all errors remain my own.

<sup>1</sup> This inchoative reading is usually left unanalyzed and mentioned together with other uses of the *imperfecto*, viz. *imperfecto* of politeness, *imperfecto* of dreams and children’s games, *imperfecto* in conditionals and possibly some others. I believe that these special uses of the *imperfecto* represent a departure from a temporal realm (i.e. they are presumably mostly modal), something that cannot be said about the inchoative reading.

<sup>2</sup> Of course, some of these types of sentences can also be rendered in the past progressive (*iba a ensayar* “was going to rehearse”) in some cases and in the conditional (*ensayaría* “would rehearse”) in some others. Differences (perhaps dialectal) in those variants, if any, are not explored here. Also, for some speakers, some verbs may sound more suitable with the periphrasis than with the *imperfecto*, a subject for another paper.

Furthermore, this inchoative sense is interesting given that it is related to telic aktionsart, as opposed to all other attested uses of the *imperfecto*, which always yield atelicity for the clause to which they belong (cf. Cipria & Roberts 2001, henceforth C&R).

### 1. *Aspect, aktionsart, and the two past tenses of Spanish*

This section assumes insights and analyses from C&R (2001) and Cipria (1996); the reader is referred to those works for finer details and extensive examples. An informal description is presented, with some formal descriptions presented as needed for the central issue at hand. I consider aspect to be a morphological category (perfective/imperfective), as expressed in the suffixes of *imperfecto* and *pretérito*. Aktionsart (atelic/telic) is a semantic notion correlated to clauses, which cannot be determined on the verb's meaning alone. Atelic aktionsart is central to our definition of the *imperfecto*. Atelicity crucially involves the subinterval property of Dowty (1987) or subsituation property of C&R, by which whatever is true for a situation is also true for its subsituations and even supersituations, displaying distributivity and cumulativity. The definition of Dowty's subinterval property is presented in (3):

#### (3) **The Subinterval Property for Atelic Aktionsarten**

If  $\delta$  is an atelic predicate, then necessarily,  $\delta(x_1, \dots, x_n)$  is true for interval  $I$  if and only if  $\delta(x_1, \dots, x_n)$  is true for all subintervals  $I'$  of  $I$ .

In the aspect/aktionsart literature, there is extensive illustration of the well-known effect on the final aktionsart of the clause of the mass/count distinction for NP objects, presence of terminal locative PPs, and durative adverbials like *for an hour/in an hour* (see section 4). In Cipria (1996) and C&R it is demonstrated that the use of the *imperfecto* brings about atelicity for the clause it is part of, regardless of the presence of telicity triggers like count-NP objects and subjects, terminal locative PPs such as *to school*, and durative adverbials like *in an hour*.<sup>3</sup> In other words, the use of the *imperfecto* overrides any

<sup>3</sup> Compare the English examples in (i) with the Spanish examples in (ii) and (iii) containing the two past tenses and same object NPs:

- |  |  |
|--|--|
| (i) a. Laura drank beer. (atelic)  | b. Laura drank a beer. (telic)   |
| (ii) a. <i>Laura tomó cerveza.</i> (atelic)<br>Laura drink <sub>PRET</sub> beer<br>"Laura drank beer."                 | b. <i>Laura tomó una cerveza.</i> (telic)<br>Laura drink <sub>PRET</sub> a beer<br>"Laura drank a beer."                 |
| (iii) a. <i>Laura tomaba cerveza.</i> (atelic)<br>Laura drink <sub>IMPF</sub> beer<br>"Laura drank/was drinking beer." | b. <i>Laura tomaba una cerveza.</i> (atelic)<br>Laura drink <sub>IMPF</sub> a beer<br>"Laura drank/was drinking a beer." |

possible telic value for the whole clause. The *pretérito*, on the other hand, is compatible with both telic and atelic aktionsart (cf. fn. 3). For the purposes of this chapter, I concentrate on the interaction of the durative adverbials with *imperfecto* and *pretérito*, which I explore in detail in section 4.

## 2. *The imperfecto*

The truth conditions for the *imperfecto* involve a core meaning: built-in atelicity, as it were, which has three main subcases associated with contextual restrictions, in addition to the temporal element PAST. The reader is referred to C&R (2001) for a complete fragment, detailing conditions and constraints, within a situation semantics framework. One can briefly describe the interpretation given to the *imperfecto* as modal in nature, involving universal quantification over situations. It has three subcases, which describe three possible types of domain restriction on the universal quantification relation; these are given in terms of permissible modal accessibility relations. The first case is a totally realistic interpretation, designed to yield the simple atelic reading of the *imperfecto*, as in (4a); the second case concerns the progressive interpretation (cf. (4b)); and the third case, the habitual (cf. (4c)), illustrates ‘characteristic subsituations.’<sup>4</sup>

- (4) a. *Ibamos a la playa.*  
 go<sub>IPLU.IMPF</sub> to the beach  
 “We went to the beach.”

---

The same effect of the *imperfecto* is observed in (iv), with a usual telicity trigger, a terminal locative PP *al colegio* ‘to school’; the English parallel is telic: *Juana took the children to school*:

- (iv) a. *Juana llevó a los chicos al colegio.*  
 Juana take<sub>PRET</sub> to the children to-the school  
 “Juana took the children to school.” (telic)
- b. *Juana llevaba los chicos al colegio.*  
 Juana take<sub>IMPF</sub> the children to-the school  
 “Juana took/was taking the children to school.” (atelic: habitual or progressive)

<sup>4</sup> For easier reference to some of the formalization, I offer the basic fragment in (i):

(i) **Meaning of the *Imperfecto*** (C & R 2001:27)

$$\| \text{IMPERF} \phi \|^{s, \text{ST}} = 1 \text{ iff } \exists s' \leq w_s [s' <_t \text{ST} \ \& \ \forall s'' [s'' \leq s' \rightarrow \forall s''' [\text{R}(s''', s'')$$

$\rightarrow \text{exemplify}(s''', \phi)]]]$ , where either:

- Totally realistic case:  $\text{R} = \{ \langle s, s' \rangle : s = s' \}$
- Progressive case:  $\text{R} = \{ \langle s, s' \rangle : s \text{ is an inertia-situation for } s' \}$ , or
- Habitual case:  $\text{R} = \{ \langle s, s' \rangle : s \text{ is a characteristic sub-situation of } s' \}$

- b. *Ibamos a la playa cuando nos encontramos con Miguel.*  
 go<sub>1PLU,IMPF</sub> to the beach when RECIPR meet<sub>1PLU,PRET</sub> with Miguel  
 “We were going to the beach when we ran into Miguel.”
- c. *Ibamos a la playa los domingos.*  
 go<sub>1PLU,IMPF</sub> to the beach the Sundays  
 “We went/used to go to the beach on Sundays.”

All other senses usually ascribed to the *imperfecto* can be derived from these three. The habitual involves a characteristic counterpart relation and the progressive is based on the idea of inertia. Informally, an inertia situation (modeled after Dowty’s 1979 inertia worlds) for a situation *s* is one that begins just like *s*, but continues in the way that *s* would continue were there no modification to the course of events as they had developed up to that point.<sup>5</sup> The relevant interpretation of the *imperfecto* for the purposes of this chapter is the futurate, which I consider a subtype of the progressive interpretation. In fact, the inchoative reading of the *imperfecto* only comes about in those situations covered by a futurate sense.

The present analysis of the *imperfecto* presents it as a multiply ambiguous form, with the common basis of atelicity, which unifies all the readings. This contrasts with de Swart’s (1998) approach, where coercion is used to derive all of these contextual possibilities, for the French *imparfait*. A fuller discussion of this contrast is presented in section 5.

### 2.1 The futurate reading

It has been widely recognized that progressives can have a futurate reading (cf. Dowty 1977, 1979, among many others). So the futurate reading is assumed here to be simply a subtype of the progressive interpretation of the *imperfecto* (cf. 4(a) with (5)):

- (5) *Hasta ayer, íbamos a la playa de vacaciones pero hoy Pepa dijo*  
 until yesterday go<sub>1PLU,IMPF</sub> to the beach on vacation but today Pepa say<sub>PRET</sub>  
*que no hay dinero para eso.*  
 that not there-is money for that  
 “Up until yesterday we were going to the beach on vacation but today Pepa said that there is no money for that.”

With this futurate reading, the intention to go to the beach holds over some past interval, and also, then, over any subinterval of that interval. That is, every

---

<sup>5</sup> Subsequent updated versions related to the ideas of inertia and branching time have been developed by Abusch (1985), Landman (1992), and Portner (1998).

subinterval of the relevant past interval verifies an instance of ‘intending to go to the beach.’

However, there is another important assumption underlying this analysis and it concerns the fact that an event consists not only of the changes in state usually associated with that type of event (cf. Dowty 1977), but also with what Moens and Steedman (1988) call a preparatory process: a subpart of the event before any culmination (of the change of state) occurs, during which the preparations for its occurrence are completed. If this assumption is made, then the truth of a clause in the *imperfecto* might be true under the progressive reading if the preparatory phase of the event (the ‘*x-ing*’) is underway. If we include the period during which one holds intentions to perform some act as part of the preparatory phase of an extended event, then the extended event is in progress during the preparatory phase, during the period when one holds those intentions. If one’s intentions are carried out as planned, then in all the inertia situations corresponding to that period, the event itself will come about, without the need to assume that the eventuality is fully realized (a common assumption about the progressive). Furthermore, the preparatory phase need not involve the intentions of a planner or an agent, but may instead simply reflect the fact that all the wheels are in motion that would ordinarily lead to an event like the sun setting, as in (6). If we take this position, then, the futurate readings are a subtype of the progressive.<sup>6</sup>

- (6) *Eran las 6. Los campesinos comenzaron a preparar el fuego. El sol se*  
 be<sub>IMPF</sub> the 6 the peasants start<sub>PRET</sub> to prepare the fire the sun 3.REFL  
*ponía a las 6:50.*  
 set<sub>IMPF</sub> at the 6:50  
 “It was 6 o’clock. The peasants started to prepare the fire. The sun was (going) to set/was setting at 6:50.”

### 3. *The inchoative reading of the imperfecto*

#### 3.1 *The characterization of the inchoative reading*

I understand the term ‘inchoative’ (also inceptive, initiative, or ingressive) to refer strictly to the beginning of a situation (Comrie 1976:19). But in the literature on aspect/aktionsart, the inchoative or inceptive terms have largely been used to indicate a change of state, as in Dowty’s (1986:50) English example:

<sup>6</sup> Of course, the ‘intention-in-the-past’ label used in C&R (2001) and in traditional works on Spanish is also covered under this assumption. Example (6) shows that there need not be a volitional planner; future events may be based on previous experience such as that provided in almanacs.

- (7) John went over the day's perplexing events once more in his mind. Suddenly, he was fast asleep.

I believe this does not correspond to a *strictly* inchoative reading, since even the use of an adverbial like 'suddenly' does not guarantee that the actual beginning of the situation is asserted; there are also some pragmatic restrictions involved.

For my purposes, I will initially assume that a strictly inchoative reading with the *imperfecto* arises (a) when there is adverbial indication (e.g. *en una hora* "in an hour") and (b) in situations that indicate a cause and effect relationship (cf. Lascarides 1992). I believe these to be the only ways in which a clear beginning of a situation can be ascertained. Note the following example, modeled after Kamp and Rohrer (1983):

- (8) *Laura encendió la lámpara. La brillante luz la encandilaba.*  
 Laura turn-on<sub>PRET</sub> the lamp the bright light her<sub>ACC</sub> blind<sub>IMPF</sub>  
 "Laura turned the lamp on. The bright light blinded her."<sup>7</sup>

It is often assumed that states may overlap a preceding event in discourse and thus are taken to include the last reference time introduced. In the case of (8), we are dealing with a process that obviously has atelic aktionsart. In this case, the turning on of the lamp causes the blinding light. The cause and effect relationship allows us to see the inception of the blinding process more clearly. But when we are dealing with states, the inception or inchoation is less obvious. Witness the following English examples, from Lascarides (1992:944):

- (9) a. Max opened the door. The room was pitch dark.  
 b. Max switched off the light. The room was pitch dark.

In relation to (9a) there is nothing in our knowledge of the world that tells us that opening a door can cause a room to become dark (unless the interpreter of the discourse stretch has information to the contrary). However, in (9b) the cause-effect relationship is clear, thus facilitating the understanding of the second sentence as asserting the beginning of a state that has just come about and, at the same time, overriding the usual effect of overlap that states display with respect to the time of the immediately preceding situation in discourse (usually associated with an event, or a telic proposition).

---

<sup>7</sup> This is their French example:

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

It is important, then, to underscore the difference between strictly inchoative readings like those in (1) and (9b) and situations that simply introduce a new state (resulting from a change of state), as in (7) and (9a), but that do not assert the beginning or the initial phase of a situation.

As hinted at earlier, there seems to be a great deal of pragmatic judgment involved in the evaluation of the data. If we compare (10) with a similar example in (11), the latter may not suggest a *strict* inchoative reading, given our perceptions about an instantaneous situation like *asesinar* “to murder” when juxtaposed with a more temporally extended situation as that indicated by *un día antes* “the day before,” which does not refer to a point-like situation as required by an act like ‘to murder.’ In other words, there is no way to know, based on the information given, when (on that particular day) the actual inception of the murder took place. In addition, as an anonymous reviewer pointed out, (11) has a passive verb phrase. This adds to the murdering event, the resulting state of having been murdered. And, obviously, the inherent homogeneity of states does not allow for a clear assessment of an inception phase; this is part of how we humans perceive reality around us and how pragmatic inferences arise. The cause-effect relationship is one example of how pragmatic inferences involving states can be canceled (cf. (9b)).

- (10) *Un cuarto de hora después dos grapos asesinaban a un policía armado.*  
 a quarter of hour later two grapos murder<sub>IMPF</sub> to a policeman armed  
 “A quarter of an hour later two members of GRAPO murdered an armed policeman.”
- (11) *Un día antes, en Santiago de Cuba, era asesinado Frank País.*  
 a day before in Santiago de Cuba be<sub>IMPF</sub> murdered Frank País  
 “The day before, in Santiago de Cuba, Frank País was murdered.”  
 (Both examples are from Butt & Benjamin 2000:212)

#### 4. *The inchoative (telic) reading vs. other available readings and the effect of adverbials*

##### 4.1 *Observed effects*

Before turning to the analysis of the crucial examples with the adverbials (*en una hora/por una hora*) in detail, recall that the *imperfecto* always brings atelic aktionsart to its clause, regardless of the presence of telicity triggers (cf. section 1). It is interesting to note that what I have been calling the inchoative reading of the *imperfecto* is the only scenario in which the atelicity of the *imperfecto* does not ‘get through,’ as it were.

The English examples in (12) show the usual behavior of these adverbial phrases; informally: *for an hour* combines with an atelic situation to give another atelic situation, that is, for the hour period, ‘Frida-rehearse-the-libretto’



is true at every subsituation of that hour period (the activity reading of the predicate). *In an hour* gives as its output a telic situation: the event of rehearsing the entire libretto took an hour or the whole situation of rehearsing the libretto was completed in an hour (the accomplishment reading of the predicate). Thus, in a telic situation it is not true to say that every subsituation verifies an instance of the same type of situation; that is, a telic like (12b) is true at the maximal interval for which the entire rehearsal took place and none of the subsituations of the one-hour rehearsal will describe the rehearsal of the entire libretto but, rather, only parts of it.

- (12) a. Frida rehearsed the libretto for an hour. (atelic)  
 b. Frida rehearsed the libretto in an hour. (telic)

The Spanish *en una hora* and *por una hora* yield similar results when combined with a *pretérito* head since this tense/aspect form is compatible with both telic and atelic clauses (cf. (13a,b) and fn. 3). With the *imperfecto*, we obtain the expected readings with *por una hora* (cf. (13c)), which combines with atelic or homogenous input; this is the case of the habitual, progressive, and futurate interpretations. As shown in (13d), these readings are also available when combined with *en una hora*, in spite of a contradiction<sup>8</sup>: *en una hora* requires a telic argument while the *imperfecto* requires that its argument be atelic (I return to this issue later). But the interesting point about 13(d) is that there arises an extra reading, which is telic in nature and would appear as a counterexample to the generalization that the *imperfecto* always brings atelic aktionsart to its clause.

- (13) a. *Frida ensayó el libretto por una hora.*  
 Frida rehearse<sub>PRET</sub> the libretto for an hour  
 "Frida rehearsed the libretto for an hour." (atelic)  
 b. *Frida ensayó el libretto en una hora.*  
 Frida rehearse<sub>PRET</sub> the libretto in an hour  
 "Frida rehearsed the libretto in an hour." (telic)  
 c. *Frida ensayaba el libretto por una hora.*  
 Frida rehearse<sub>IMPF</sub> the libretto for an hour  
 "Frida used to rehearse/was rehearsing the libretto for an hour."  
 (atelic: habitual or progressive)

---

<sup>8</sup> Again, see C&R (2001) for ample discussion and formalization.

- d. *Frida ensayaba el libreto en una hora.* (shown earlier as (1))  
 Frida rehearse<sub>IMPF</sub> the libretto in an hour  
 “Frida rehearsed/used to rehearse/was rehearsing/intended/was to rehearse the libretto in an hour.”  
 (telic: inchoative; or atelic: habitual, progressive or futurate)

For ease of reference, I include in (14) the definition of telicity, according to Dowty (1987):

- (14) If  $\delta$  is a telic predicate, then the truth of  $\delta(x_1, \dots, x_n)$  for interval  $I$  entails that  $\delta(x_1, \dots, x_n)$  is false for all proper subintervals  $I'$  of  $I$ .

As expressed earlier, the telic interpretation described for (13d) might appear to be a counterexample to the generalization that the *imperfecto* always yields atelic aktionsart. However, this reading does not share the truth conditions available for (13b), where the entire rehearsal took one hour. Instead, the meaning is that in one hour the rehearsal began/was to begin; the simple telic change of state is from one in which no rehearsal was underway to one in which it had begun.<sup>9</sup> In general, when a telic adverbial like *en una hora* “in an hour” occurs with an atelic clause, one way of making the result felicitous is to shift to an inchoative interpretation, where the endpoint of the hour period marks the beginning of the process or state corresponding to the atelic clause—here, the process of rehearsal. We can see this in English if we give *in an hour* wide scope over the progressive, which is always atelic:

- (15) Frida was rehearsing the libretto in an hour. (telic: inchoative)

Sentence (15), then, has the following rough syntactic structure:

- (15') **PAST**[*in an hour*[**PROG**[Frida-rehearse-the-libretto]]]

It should be pointed out that *in an hour/for an hour* are considered VP modifiers, so-called ‘aspectual adverbials’ (cf. Dowty 1979), and the progressive is also considered a VP modifier. Tenses (past/present) are considered sentential modifiers.

Following C&R (2001), the other (atelic) readings available for (13d) may be accounted for by assuming that in them the adverbial *en una hora* takes as its argument the tenseless/aspectless *ensayar el libreto*, which is indeterminate

---

<sup>9</sup> Scope issues and other problems aside, the analysis of this state of affairs could include an element along the lines of the BECOME operator, a well-known proposal by Dowty (1979).

with respect to aktionsart,<sup>10</sup> like its English counterpart; the *imperfecto* then applies to the resulting clause to yield atelic aktionsart. Example (13d), for the atelic readings, would then have the following syntactic structure:

(13d') **IMPF**[*en una hora*[Frida-rehearse-the-libretto]] (atelic readings)

In the habitual reading of 13(d), we get the interpretation that it was a habit of Frida's to rehearse the libretto in an hour and that there is a period in the past where all its characteristic subsituations exemplify Frida's taking exactly one hour to rehearse the libretto. In the progressive reading of (13d) we get the simple atelic reading that Frida was in the process of rehearsing the libretto in an hour. The context in (16) brings out the futurate reading of (13d):

(16) *Mañana, Frida ensayaba el libreto en una hora y después se*  
 tomorrow Frida rehearse<sub>IMPF</sub> the libretto in an hour and then REFL  
*encontraba en el cine con Salma.*  
*pro-meet<sub>IMPF</sub> at the movie-theater with Salma.*  
 "Tomorrow, Frida was rehearsing/was to rehearse the libretto in an hour and later she was meeting/was to meet Salma at the movie theater."

*Mañana* "tomorrow" brings out the interpretation that Frida was planning to rehearse the libretto in an hour, expecting that it will take an hour to do that and that after that, she would do something else. As per the assumption that the futurate is a subtype of the progressive case of the *imperfecto* and the additional assumption of a preparatory process, it is true for some past period that at every relevant subsituation, it was true that Frida was planning to rehearse the libretto in an hour.

#### 4.2 Accounting for the telic reading of (13d)

Consider the sentence in (17), with *en una hora* in focus position:

(17) **En una hora**, *Frida ensayaba el libreto.*  
 in an hour Frida rehearse<sub>IMPF</sub> the libretto  
 "In an hour Frida was rehearsing/intended to rehearse/was to rehearse the libretto."

If we give *en una hora* a focus position (which I assume corresponds to the sentential modifier position logically), we do not get any of the other readings, such as the habitual or plain progressive reading; rather, the futurate reading is the only interpretation possible, indicating, again, that the rehearsal was to

<sup>10</sup> Compare earlier discussion of the activity/process versus the accomplishment reading of the predicate.

begin in an hour from the past utterance time, asserting the beginning point of the rehearsal, or the inchoative interpretation. Recall the scoping parallel with the English progressive, which yields a telic reading when the adverbial has wide scope over the progressive.

C&R (2001) offer an explanation of the telic reading of (13d) that involves the ambiguity of the adverbial. In the telic reading, *en una hora* does not have a durative meaning but, rather, a referential meaning of merely locating some event in time, parallel to some adverbial like ‘soon.’ Thus, as stated earlier, this reference time adverbial temporally anchors (and asserts) the beginning of the hour period at which the rehearsal of the libretto was to begin, hence the inchoative interpretation. Those authors hint at a solution along the lines of an internal adverbial treatment (cf. Dowty 1979), whereby *en una hora* is internal with respect to the tense portion of the *imperfecto* but has wide scope over the aspectual portion (this becomes necessary given that the *imperfecto* combines tense and aspect in one morphological form). A rough representation would look like this:

(13d'') PAST [*en una hora*[ASP-IMPF[Frida-rehearse-the-libretto]]] (telic reading)

In this representation, I assume that *en una hora* is a sentential modifier, corresponding to the focus situation argued for earlier.

Although the internal adverbial proposal and the interpretation of the adverbial as a reference time adverbial explain the difference in readings, there is still more to be said about adverbial scope and why it allows for a telic reading, which, as mentioned earlier, is the only case (based on the available data) where the atelicity of the *imperfecto* does not yield atelic aktionsart for its clause. In other words, in the case of count-NP complements and terminal locative PPs (usual triggers of telicity), the *imperfecto* overrides their effect to entail atelicity for the whole clause (cf. fn. 3).

Dowty’s preface to the new edition of his book (1991:xxiv) mentions this particular problem of the scope of adverbials and of the difference in readings when the adverbial is fronted; but, so far, no solution that I am aware of has been offered in the literature for that problem.

An explanation of the mismatch between a telic adverbial and an atelic (*imperfecto*) clause can be couched in terms of coercion. That is, the mismatch in telicity value causes a shift of interpretation (from atelic to telic) through the process of coercion, in this case, and inchoative coercion, as in Moens and Steedman (1988). I review coercion effects in section 5. Coercion does explain how contextual reinterpretation takes place when there is a contradiction or a clash of requirements; that is, in the present case it can explain why the

inchoative readings in (6), (10) and (13d) are possible. However, this still does not explain why it is the case that adverbials like *en una hora* are like barriers in that they block the atelicity-inducing effect of the *imperfecto*. In spite of this proviso, the inchoative reading of the *imperfecto* is a nice addition to the increasing body of literature on coercion.

Before turning to the discussion of coercion, though, an anonymous reviewer pointed out that there are nontrivial differences between the scope configuration suggested earlier and the analysis of perfective/imperfective with adverbials in de Swart (1998). In what follows, I try to take a first stab at the issue, although I am not able to develop it fully, given space limitations. In fact, the account presented here presents aspect as an independent category, its status being important for how it relates, in terms of scope, with adverbials like *en una hora*. De Swart views the *imparfait* and *passé simple* as aspectually sensitive tense operators. This predicts, according to her, that perfective and imperfective are not independent categories and are not expressed by independent morphology, and that a distinction only exists in the past domain. In de Swart's account, an inchoative operator ( $C_{he}$ ) presumably has narrow scope with respect to the PAST operator (which is tense and aspect in one) and an adverbial like 'in an hour' would be internal to the inchoative operator (i.e. it has narrow scope with respect to the operator). In the present account, *en una hora* has narrow scope with respect to the tense portion of the *imperfecto* but wide scope over the aspect portion of it.

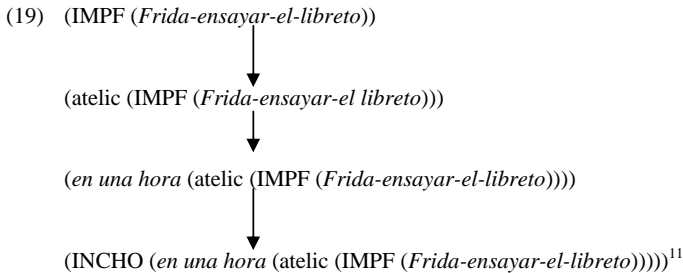
### 5. *Coercion*

Another way of looking at how verbal clauses combine adverbials and other elements in order to yield different aspectual or aktionsart values is to think of those combinations as transitions in a network, as in Moens and Steedman (1988). When the input to a combinatorial operation or a transition does not match the specifications to yield a given reading, then a process of reinterpretation takes place. Moens and Steedman call a process like this 'coercion.' In this framework, aspectual auxiliaries, tense morphemes, adverbials and other elements are seen as functions (operators in the subsequent literature) that can *coerce* their inputs (or arguments) to the appropriate type. Thus, witness one of their examples in (18):

- (18) Sandra was hiccuping.

In order to arrive at the progressive form of (18), Moens and Steedman show that first, the 'point proposition' (their term for a type of telic event like *Sandra hiccuped*) is coerced (by the progressive) into a process of iteration of the

event, thereby rendering a ‘progressive state’ viewed as ongoing. Thus, the transition by means of coercion is from a telic type to an atelic type. Later, I discuss my views on a possible coercion treatment for all of the attested readings of the *imperfecto*. However, I believe that the clearest candidate for this type of approach is the telic reading of (13d), thus observe (19), which presents a sketch of what the coercion process might look like (with the terminology used in this chapter):



In Moens and Steedman (1988), arches and arrows are used to represent the transitions from one type to another. In (19), then, first the IMPF has applied to the atomic predicate *ensayar-el-libreto* “rehearse-the libretto” to yield an atelic proposition. Recall that I have accounted for this by the effect of the very meaning of the *imperfecto*, which features the central element of the subinterval or subsituation property, and the different readings are accounted for by a modal accessibility relationship associated with the different contextual restrictions. However, for our purposes, the second transition is crucial: the adverbial *en una hora* “in an hour” coerces the atelic proposition into a telic one and then it receives an inchoative reading.

De Swart (1998) uses the concept of coercion to account for the different readings of the French *imparfait*, which, in the present approach, are accounted for by contextualized variants (subcases) of one core meaning of the *imperfecto*, assuming *imparfait* and *imperfecto* behave similarly. De Swart utilizes coercion operators to account for habitual, progressive, iterative, and inchoative readings (the latter is not described in much detail). For de Swart, tense morphemes on the *imparfait* and *passé simple* do not contribute aspectuality but merely locate an ‘eventuality description’ (her ontological

<sup>11</sup> Here I use INCHO as a coercion function or operator, which does not actually appear as such in Moens and Steedman (1988). So ‘INCHO’ here is my cover term for Moens and Steedman’s original concept coupled with de Swart’s (1998) inchoative variant of a coercion superoperator.

label) in time. These two past tenses of French are “aspectually-sensitive tenses,” according to de Swart (1998, 2000). The difference between these two French past tenses lies in the way they select for the arguments with which they combine: the *imparfait* calls for an atelic eventuality description while the *passé simple* calls for a telic eventuality description. Coercion is called for whenever there is a mismatch in value between a function (or modifier) and an argument with which it may potentially combine. Note that, in my account, the *prétérito* (which, I assume, is parallel to the *passé simple*), is compatible with both telic and atelic aktionsart, that is, the *prétérito* ‘passes on’ the telicity value of the VP it combines with (cf. examples (iia, b)).

A different view of coercion is offered in Bonami (2002), who considers implicit aspectual operators for the *imparfait*, which are lexically licensed. The proposal is couched in a Head-Driven Phrase Structure (HPSG) framework, enriched with minimal recursion semantics and some semantic assumptions from de Swart (1998, 2000). Very roughly, Bonami seems to argue, in relation to the French *imparfait*, that there is no need of a clash or contradiction for habitual or iterative readings<sup>12</sup> to arise. That is to say, according to him, it is not necessary to argue that there have been coercion operators mediating those interpretations but that, instead, they can be *lexically* licensed by an adverbial, or by the mere *imparfait* progressive operator. In other words, he assumes that there are implicit aspectual operators but that they are not coercion operators. In this respect, he argues that de Swart’s analysis overgenerates coercion processes and it undergenerates in some other cases.

Recall that, in my treatment of the *imperfecto*, the different attested readings arise from the truth conditions of the *imperfecto* itself, with the subinterval or subsituation property being the unifying element for all the subcases. It seems, then, intuitively accurate that there should not be a coercion operator for every single subcase of the *imperfecto*, that is, the plain atelic, the progressive, and the habitual (other readings can be derived from these as was shown earlier). Furthermore, if we considered every contextual variant of the *imperfecto* to be licensed by coercion operators (as in de Swart 1998), the special shifted reading we get for the inchoative, that is, a coerced reading, would not be such an odd occurrence, as in fact seems to be the case. Examining the issue from another point of view: if we assumed that every time the *imperfecto* overrides the effect of telicity triggers (cf. fn. 3 and atelic readings of (13d)) we are dealing with coercion, then we would still have to deal with the special telic reading of (13d) and explain why it is not the case that the use of the *imperfecto* yields atelicity for the whole clause.

---

<sup>12</sup> Iterative readings are discussed in C&R (2001:36-38).

This inchoative example, I believe, is the clearest candidate for coercion, since the truth conditions of the *imperfecto* (specifically its subinterval property or atelicity) do not seem to be able to apply. Or stated differently, the transition from atelic to telic does not arise even in the presence of usual telicity triggers like count NPs or terminal locative PPs (cf. examples (i)-(iii)). In this case, a lexicalist approach, similar to Bonami's (2002) seems to be more suited to explaining these diverse results, in addition to the C&R (2001) analysis of different contextual restrictions on the core atelic meaning of the *imperfecto*. In the case of the inchoative reading of the *imperfecto*, surrounding sentences in the context and adverbials (in the appropriate scopal configuration) do coerce a special telic case. Recall that usual telicity licensors like count NPs or terminal locative PPs do not 'license' or coerce a telic reading for the *imperfecto* (as in (i)-(iii)); that is, they never override the inherent atelicity of the *imperfecto*, while adverbs and surrounding discourse do.

In summary, coercion is an undeniable phenomenon; however, it does not need to be invoked every time one and the same form can give different readings. That is, *imperfecto* clauses keep being atelic no matter what interpretation we give them (i.e. habitual, progressive, and so on). But in the case of the inchoative reading, there has been a major change of aktionsart type (from atelic to telic), which is why this is a clearer case of coercion. It is also a case of coercion because this does not happen in any other context, where the inherent atelicity remains. In other words, in the present account, we capture the intuition that the *imperfecto*'s inherent atelicity allows for contextual ambiguities but the core of the tense, its meaning, remains unchanged (except in the inchoative reading) and there is no need to have the *imperfecto* forced in some way in order to have three different meanings that do have something in common.

If we assume that coercion applies across the board and that it is a more or less natural phenomenon whenever there is a type mismatch, then we would still need to account for the *pretérito*'s compatibility with both telic and atelic aktionsart, in which case there is no need for any kind of coercion. It seems counterintuitive, then, to consider only the *imperfecto* as a candidate for coercion.

While more discussion cannot be included due to space restrictions, perhaps a combination of my approach and terminology with Bonami's (2002) would account elegantly for facts in Spanish and French. For example, we can say that the strict inchoative reading is 'lexically coerced' by specific adverbs



or surrounding discourse. It also remains to be seen whether there is a type of clear coercion (as in the telic inchoative) with the *pretérito*.

## 6. Conclusion

I have examined the interesting case of the inchoative interpretation of the *imperfecto* by arguing (with C&R 2001) that this is a shifted reading resulting from coercion and from the interpretation of accompanying adverbials such as *en una hora* as reference time adverbials. I have also clarified the difference between what was termed a 'strict' inchoative reading and a reading that reflects a change of state (though not necessarily the onset of the new state). My examination highlighted the fact that, based on available data, there is a highly important difference in how the atelicity of the *imperfecto* affects the aktionsart of its clause in different environments, viz. mass-/count-NP objects, terminal locative PPs, and temporal adverbials. I also explored the notion of coercion as handled by de Swart (1998, 2000) and Bonami (2002), and concluded that, as far I have been able to ascertain, such a process is only justifiable for the special telic reading of the *imperfecto* arising in combination with *en una hora*-type adverbials.

With this chapter, I hope to have opened up a forum for discussion on a special interpretation of the *imperfecto* that, as far as I know, has been neglected in the formal literature on Spanish. The issues touched upon here may also shed light on the behavior of the *imperfecto* in discourse. In turn, the discussion has brought up a mere tip of the iceberg on the issue of temporal adverbials in general, and of adverbial scope in particular, which only further research will help elucidate.

## REFERENCES

- Abusch, Dorit. 1985. *On Verbs and Time*. Ph.D. dissertation, University of Massachusetts at Amherst.
- Bonami, Olivier. 2002. "A Syntax-Semantics Interface for Tense and Aspect in French". *Proceedings of the 8<sup>th</sup> International HPSG Conference, Norwegian University of Science and Technology, 3-5 August, 2001* ed. by Frank van Eynde, Lars Hellan, & Dorothee Beer-mann, 31-50. Stanford, CA: CSLI Publications.
- Butt, John & Carmen Benjamin. 2000. *A New Reference Grammar of Modern Spanish*. Chicago: McGraw-Hill.
- Cipria, Alicia. 1996. *The Interpretation of Tense in Spanish Complement Clauses*. Ph.D. dissertation, The Ohio State University, Columbus.
- Cipria, Alicia & Craige Roberts. 2001. "Spanish *Imperfecto* and *Pretérito*: Truth conditions and aktionsart effects in a situation semantics". *Natural Language Semantics* 8.297-347.

- Comrie, Bernard. 1976. *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.
- Dowty, David. 1977. "Toward a Semantic Analysis of Verb Aspect and the English 'Imperfective Progressive'". *Linguistics and Philosophy* 1:1.45-78.
- , 1979. *Word Meaning and Montague Grammar*. Dordrecht: Kluwer.
- , 1986. "The Effects of Aspectual Class on the Temporal Structure of Discourse: Semantics or pragmatics?". *Linguistics and Philosophy* 9:1.37-62
- , 1987. "Aspect and Aktionsart". Ms., The Ohio State University, Columbus.
- , 1991. *Word Meaning and Montague Grammar*. 2nd ed. Dordrecht: Kluwer.
- Kamp, Hans & Christian Rohrer. 1983. "Tense in Texts". *Meaning, Use and Interpretation of Language* ed. by R. Bäuerle, C. Schwarze, & A. von Stechow, 250–269. Berlin: de Gruyter.
- Landman, Fred. 1992. "The Progressive". *Natural Language Semantics* 1.1-32.
- Lascarides, Alex. 1992. "Knowledge, Causality and Temporal Representation". *Linguistics* 30.941-973.
- Moens, Marc & Mark Steedman. 1988. "Temporal Ontology and Temporal Reference". *Computational Linguistics* 14:2.15–28.
- Portner, Paul. 1998. "The Progressive in Modal Semantics". *Language* 74.760-787.
- de Swart, Henriette. 1998. "Aspect Shift and Coercion". *Natural Language and Linguistic Theory* 16.347-385.
- , 2000. "Tense, Aspect and Coercion in a Cross-Linguistic Perspective". *Proceedings of the Berkeley Formal Grammar Conference* ed. by Miriam Butt & Tracy Holloway King, 1-20. Berkeley, CA: CSLI Publications.



**REINTERPRETING THE CV TRANSITION\***  
**EMERGENCE OF THE GLIDE AS AN ALLOPHONE OF THE PALATAL LATERAL**

LAURA COLANTONI  
*University of Toronto*

**0. Introduction**

Palatal laterals have been part of the inventory of most Romance languages. At present, however, different rates of maintenance are observed, from almost complete loss in French (Nyrop 1923) to preservation in Portuguese (Silva 1999) and Catalan (Recasens et al. 1993; Recasens & Pallarès 2001). Italian (Bladon & Carbonaro 1978) and Spanish exhibit a variable rate of maintenance across varieties. It has been documented that in Spanish dialects palatal laterals have been (Quilis 1993; Lipski 1994, among others) or are being substituted by either a palatal glide [j] or a palatal fricative [ç] (De los Heros 1997).

In Argentine Spanish, palatal laterals have been previously reported in the northwest (provinces of San Juan and Catamarca), and in the northeast (provinces of Corrientes, Chaco, Formosa, and Misiones) (Vidal de Battini 1964).<sup>1</sup> In a recent study (Colantoni 2001), however, palatal laterals were not observed in San Juan, and different rates of maintenance were reported for the province of Corrientes, ranging from 100% in the north in 50% to the south. In the locations with a variable rate of maintenance, a glide emerges as the preferred allophone of the palatal lateral. Both the palatal lateral and the glide maintain an opposition with a voiced palatal affricate [dʝ], as in (1).

(1) <i>ca</i> [λ]ó <i>calló</i> “he/she stopped talking”	vs.	<i>ca</i> [dʝ]ó <i>cayó</i> “he/she fell down”
--	-----	--

Example (1) illustrates that the phonetic realization of the phonological opposition in Corrientes Spanish is different from most Spanish dialects, in

---

\* I want to express my gratitude to the three anonymous reviewers for invaluable suggestions and comments. All errors remain my own.

<sup>1</sup> It has been claimed (Penny 2000) that palatal laterals are maintained due to contact with Native American languages, such as Quechua or Guarani, which have palatal laterals in their inventories (Maddieson 1984).

which the palatal lateral contrasts with a palatal fricative or glide. Thus, when the palatal lateral disappears from those dialects and is replaced by a glide, a merger immediately takes place in a phenomenon known as *yeísmo*.<sup>2</sup> In contrast, even if the palatal lateral disappears from Corrientes Spanish, the glide can still establish an opposition with the palatal affricate, from which it differs in articulatory and acoustic aspects. This difference, then, should allow us to analyze the evolution of the palatal lateral independently of the merger.

Thus, the present chapter focuses on the process by which the glide emerges as an allophone of the palatal lateral, and eventually replaces it. Comments about the similarities between palatal laterals and glides can be found in the literature on Spanish phonology (Lipski 1989) and dialectology (Quilis 1993), and also in studies on other Romance languages (Silva 1999). Phonological accounts of the change have been proposed within autosegmental (Lipski) and Optimality Theory (Colantoni 2001). In spite of their formal differences, both accounts assume that the result of the change is the emergence of a different segment.<sup>3</sup> I show here that the glide was indeed present before the change occurred in the CV transition. I argue, in particular, that this change could be interpreted in the spirit of Ohala's (1989) hidden variation theory, which claims that synchronic variation is a precondition for sound change. Sound change takes place when the listener, in the acquisition process, reinterprets the acoustic cues that are already present in the signal; that is, a secondary cue becomes the relevant cue to parse the segment. In the specific case of Argentine Spanish, a factor that is already present in the signal, namely, the glide-like CV transition, would be interpreted by the listener as the relevant cue of the segment, motivating the substitution of [ʎ] by the glide.

The presence of a glide-like CV transition would not account, however, for the loss of the lateral. Two alternative hypotheses should be considered. The first is that the glide becomes longer, and the palatal shorter, until the lateral eventually disappears. The alternative hypothesis is that the constriction of the lateral becomes wider, and then more similar to the glide. Based on the results obtained, I argue that this second hypothesis is the most likely explanation for the change observed.

---

<sup>2</sup> Although the most frequent realization is a palatal fricative (Navarro Tomás 1970), I assume, based on previous studies (Whitley 1995; Lloyd 1993), and on spectrographic evidence from Argentine Spanish (Colantoni 2001) that the palatal fricative is the result of a strengthening process that applies to an underlying glide.

<sup>3</sup> Although the palatal lateral and glide share most of the feature specification, the change from the former to the latter implies a qualitative difference. The glide was not present before the change took place.

Finally, this study has implications for the phonetics and phonology of palatal consonants. If there is a glide-like segment, it could be interpreted either as part of the palatal consonant, thus supporting the analysis of palatals as segments involving two articulators (Keating 1988), or as a consequence of the coarticulation between the palatal and the vowel, thus providing evidence against the two-articulator hypothesis (Recasens 1990; Recasens et al. 1993). If palatals are not phonetically complex, phonological representations that treat them as complex segments should also be revisited (Carreira 1988; Lipski 1989).

The chapter will be organized as follows: After reviewing the articulatory and acoustic characteristics of palatal laterals and glides, I present the methodology. In section 4, I analyze the palatal lateral and the transitions. Then, I account for the change in palatals and discuss the implications of the results for phonetic and phonological analyses of palatals.

## 1. *Palatal laterals and glides: Articulatory and acoustic characteristics*

### 1.1 *Palatal laterals*

The articulatory description of palatal consonants has been the object of a controversy that basically involves the determination of the active articulator. Two positions can be identified: those who argue that both the tongue blade and body act as the active articulator (Keating 1988), and those who consider that the tongue dorsum is the only articulator activated (Recasens 1990; Recasens et al. 1993). According to the first position, palatals constitute one class, and can be classified as complex segments, given the simultaneous activation of two parts of the tongue. The second position argues that, since only the tongue dorsum is activated, palatals are not complex segments. In addition, this position distinguishes alveopalatals (including [ʎ]) from true palatals, such as [j], where the dorsum-mediodorsum of the tongue is the active articulator.

The aforementioned studies do not, however, explicitly describe the Spanish palatal lateral. Keating (1988) focuses on the palatal nasal, while Recasens (1990) and Recasens et al. (1993) analyze the palatal lateral in Catalan and Italian. Palatographic evidence presented by Navarro Tomás (1970:133) suggests that there are no significant differences in the articulation of Catalan and Peninsular Spanish palatal laterals.<sup>4</sup> According to Recasens, the main active articulator is the dorsum with some possible involvement of the

---

<sup>4</sup> Although the speakers analyzed by Recasens et al. (1993) show an opening on both sides of the tongue, laterals can be also articulated with an opening on only one side (cf. Navarro Tomás 1970).

mediodorsum, and less frequently the lamina and postdorsum. The passive articulator targeted is the postalveolar and prepalatal area, with optional (and rare) constrictions at the mediopalatal, alveolar, and postpalatal zones (Recasens 1990:272). The degree of contact may vary according to the speaker and the following vowel, but only for the alveolar and mediopalatal and postpalatal zones. In those areas, the degree of contact is higher with high vowels than when adjacent to [a]. Nonsignificant differences have been observed, however, for front and back vowels in the same height dimension.

Lateral segments can be acoustically characterized by the presence of clearly defined formant patterns. Their first formant (F1) is relatively low frequency (300-400 Hz), the second formant (F2) varies according to the location of the constriction, and the third formant (F3) is relatively high in frequency. Fant (1960) indicates that F1 values correlate with the degree of opening (high F1 = more open articulation); Narayanan, Alwan, and Haker (1997:1074) conclude that F2 “can be associated with the half-wavelength resonance of the back cavity” (i.e., a longer cavity behind the constriction corresponds to a lower F2).

Two resonators are involved in the production of laterals: (a) the one formed by the opening on both sides of the tongue and (b) the cavity on top of the tongue. The presence of this latter cavity introduces an antiformant in the spectrum in the range of F2-F4 (Fant 1968). In addition, lateral segments can be distinguished from the following vowel by an abrupt change in formant transitions, especially for apical laterals.

Acoustic descriptions of palatal laterals are not abundant, probably because they are relatively uncommon cross-linguistically (Maddieson 1984). General descriptions (Ladefoged & Maddieson 1996) suggest that an F1 below 400Hz characterizes the segment. Previous studies on Romance languages (Bladon & Carbonaro 1978; Quilis et al. 1979; Silva 1999) differ in the selection of the independent variables, in the methods used to measure them, and in their sample sizes. The number of speakers included in the studies varies: eight for Spanish (Quilis et al.), two for Italian (Bladon & Carbonaro), and only one Brazilian speaker for Portuguese (Silva).<sup>5</sup> All the studies report values on the duration of the segment and indicate some characteristics of its quality. Regarding the former, there is a consensus that palatal laterals are long segments, that is, they are longer than other consonants, and, in particular, longer than other liquids (Bladon & Carbonaro; Silva). The Italian lateral may

---

<sup>5</sup> Quilis et al. (1979) analyzed the speech of seven Peninsular Spanish speakers and one Colombian. Given the unbalanced number of speakers and that results are not discriminated by speaker, I will consider these findings representative of Peninsular Spanish.

be the longest, with a duration ranging between 197 and 217 ms (Bladon & Carbonaro 1978:46), while Spanish has the shortest (79 ms in stressed syllables and 64 ms in unstressed syllables; Quilis et al.).<sup>6</sup> Intermediate values have been reported for Brazilian Portuguese (101 ms in word-initial position vs. 78 ms in intervocalic position; Silva 1999:57).

Regarding the quality of the lateral, all the studies report values of F2, which vary from 2273-2666 Hz in Italian (Bladon & Carbonaro 1978), to 2000 Hz in Spanish (Quilis et al. 1979), and 1870 Hz in Portuguese (Silva 1999).<sup>7</sup> The values of F1 and F3 are only reported on for Spanish (Quilis et al.) and Portuguese (Silva). In Spanish, F1 mean values are approximately 290Hz, while F3 ranges from 2540 Hz to 2766 Hz. In Brazilian Portuguese similar values for F1 (300Hz), but slightly higher values for F3 (around 2900 Hz), have been reported. Although some variation under different stress conditions has been observed, researchers indicate that the quality of the lateral is not significantly affected by the vocalic context (Bladon & Carbonaro; Quilis et al.).

CV transitions have not received the same attention. The variables studied in the literature include direction and duration of the transitions (Quilis et al. 1979; Colantoni 2001). Transition duration is a problematic variable, since measurements are not consistent from researcher to researcher. Values reported in the literature indicate that the mean duration in Peninsular Spanish (Quilis et al.) is 30 ms, which is half the duration observed in Brazilian Portuguese (Silva 1999). Only Silva reports formant values of the transition, but, unfortunately, the palatal lateral is only studied in one vocalic environment (i.e., preceded and followed by [a]). Thus, in order to arrive at more definitive conclusions regarding the nature of the transition, and then test the hypothesis of the presence of the glide, we have to rely on articulatory and acoustic descriptions of the palatal glide.

### 1.2 *Palatal glides*

The articulatory description of glides poses some problems (Recasens 1990), since they lack a steady-state portion. Some studies group glides with other palatals (Keating 1988), while others include them in a separate class (Recasens). The general consensus is that the palatal glide mainly differs from other palatals in the degree of the constriction (Keating), and Recasens argues that a different part of the tongue is actively involved in glides, namely the

---

<sup>6</sup> The values reported include the transitions.

<sup>7</sup> Results reported by Bladon and Carbonaro (1978) are in mels, and were converted into Hertz to facilitate the comparison.



mediodorsum, with occasional involvement of the laminal portion. According to Recasens et al. (1993:229), the tongue moves as a whole in the articulation of the glide.

Acoustic characterizations of the glide face the same problems as articulatory ones; that is, given the lack of a steady-state portion, there has been much discussion about the relevant variables that should be measured (cf. Aguilar 1997, 1999). Studies on Spanish (Borzone de Manrique 1976; Aguilar 1997, 1999) generally agree that palatal glides have more centralized formant values than the high front vowel [i], and exhibit a high degree of contextual variability. In particular, Aguilar (1997:187; cf. also Borzone de Manrique 1979) concludes that glides are very ‘permeable’ to influences of the vocalic context.

Two sources provide acoustic measurements of the quality of the Argentine Spanish glide: Borzone de Manrique (1976, 1979) reports data of Buenos Aires Spanish in the diphthong [je], while Colantoni (2001) characterizes the glide in onset position, which is found in several Argentine provinces (Córdoba, San Luis, and San Juan). Borzone de Manrique finds that the first three formants of the glides are within the following ranges, respectively: 254-471 Hz, 2246-2757 Hz, and 2756-3615 Hz.<sup>8</sup> Colantoni reports mean values that are within the range reported for Buenos Aires for the F1 (390 Hz), but lower for the other two formants (1474 and 2174 Hz, respectively).

## 2. *Palatal laterals: Phonotactic constraints*

Palatal consonants in Spanish are subject to specific phonotactic constraints. They are only attested in syllable onset position; are the only consonants disallowed in codas; cannot be followed by a front glide; cannot be preceded by [r], [n], or [l]; and are anomalous with respect to stress assignment (Carreira 1988; Lipski 1989). In particular, palatal laterals and high vowels exhibit additional combinatory constraints, a factor that is tangentially mentioned by Quilis et al. (1979), but has not been studied in detail. First, the sequence [ʎi] is not allowed in word-initial position, and word internally is almost exclusively permitted in stressed syllables.<sup>9</sup> The sequence [ʎu] is also subject to some restrictions. It is attested in word-initial position, although it is more frequent in stressed syllables, and is not observed word internally in post-tonic syllables.

---

<sup>8</sup> Borzone de Manrique (1976, 1979) reports her results in mels and in graphics. The values included were obtained from the graphics and converted into Hertz.

<sup>9</sup> The only possible exception to this generalization is found in the word *gallinero* ‘henhouse.’

The constraint against the combination of palatal laterals and the high front vowel, in particular, can be accounted for in acoustic and auditory terms, as a restriction against two similar segments. Therefore, constraints on the sequence [ki] could be interpreted as a way to avoid an obligatory contour principle violation (McCarthy 1986).

The specific phonotactic constraints to which palatals are subject have motivated their analysis as phonologically complex segments. Complexity, however, is understood in two different ways: as a sequence of two segments (Carreira 1988) or as a segment specified for two place nodes (Lipski 1989).<sup>10</sup> Carreira (1988:73) argues that Spanish palatals can be represented as [+coronal, -continuant segments] followed by a [+high, -consonantal] segment. On the other hand, Lipski proposes a representation that includes two place nodes (coronal and dorsal) associated with one time unit, and bases his analysis on Keating's (1988) findings on the articulatory characteristics of palatal segments. While the analyses differ in the representation of palatals (see section 6 for a discussion), both reinforce the conclusions of articulatory studies suggesting that palatals have a special status and, thus, that acoustic characteristics of palatals deserve a closer look.

### 3. *Methods*

Data analyzed here were taken from the linguistic-anthropologic atlas of Argentina, directed by O. Kovacci (1987). The interviews were approximately 2 hours long, and included vocabulary-elicitation tasks and narratives. According to the criteria established by the project, two speakers (a man and a woman) were selected in every location. Those speakers were lifelong residents of the location under study, had no secondary education, and their ages ranged from 25 to 65. Recording took place in a quiet place, using a portable tape recorder, a metal tape, and a unidirectional microphone.

The present chapter analyzes the speech of eight natives of four locations in Corrientes (Bella Vista, Beron de Astrada, Alvear, and San Miguel). They were selected from a larger pool of subjects (cf. Colantoni 2001) based on their high rate of maintenance of the palatal lateral. Since the goal is to describe the palatal lateral in order to account for the sound change observed in Corrientes and in other Spanish varieties, it is important to start by characterizing the segment where the rate of maintenance is the highest. Results obtained should

---

<sup>10</sup>According to Carreira (1988), palatal segments arise through a combination of palatalization and resyllabification rules. They come from an underlying sequence of a [+anterior] coronal in the coda followed by a glide in the onset of the following syllable.

be compared in the future with the realization of palatal laterals among speakers with a more advanced degree of loss of the lateral.

All the words exhibiting the palatal lateral were extracted from the signal, which was digitized at 22,000 Hz/16bits. The presence of the lateral was determined not only auditorily, but also using both the spectrogram and an FFT spectrum. The total number of words extracted was 347, of which only 270 were acoustically analyzed, due to insufficient quality of some of the recordings. The number of tokens is not equally divided among all the speakers (see appendix) for two reasons: (a) Some speakers responded to the questionnaire using lexical items that did not contain the palatal lateral; (b) some tokens had to be discarded due to the presence of background noise, coughing, and so on.

On the tokens selected, the same measurements (i.e., frequency of the first three formants and duration of the segment) were taken for the consonant and the transition, using Praat 4.0.41. Formant values were measured at the midpoint of the consonant and of the transition, which is defined as the portion that extends from the release of the lateral to the steady state of the vowel (see Figure 1). Formant values provide two pieces of information: (a) They indicate the degree and place of constriction of the lateral and (b) they offer crucial evidence to test the main hypothesis of the chapter (the presence of a glide-like CV transition). Duration measurements allow testing of two alternative hypotheses, that is, whether the change observed in Argentine Spanish involves a reduction of the closure of the lateral and a lengthening of the transition, or a change in the degree of the constriction in the closure of the lateral. Quality and duration of the following vowel were also measured, but will not be reported here, since the focus is on the characterization of the lateral segment.

Extralinguistic and linguistic variables were included in the analysis. The former involve 'location' and 'gender,' while the latter consist of 'following vowel' [a, e, o] and 'stress' (stressed vs. unstressed syllables).<sup>11</sup> The position in the word was not taken into account, since only 11 tokens were found in word-initial position. All the values were exported to Excel; statistics were calculated with SAS 8.2. Statistical significance was defined as  $p < 0.05$ .

---

<sup>11</sup> Given the phonotactic constraints restricting the sequences of palatal laterals and high vowels, words containing them were excluded when measuring the effect of the following vowel on the quality of the lateral or the transition. In the corpus, only 18 words containing those vowels were observed.

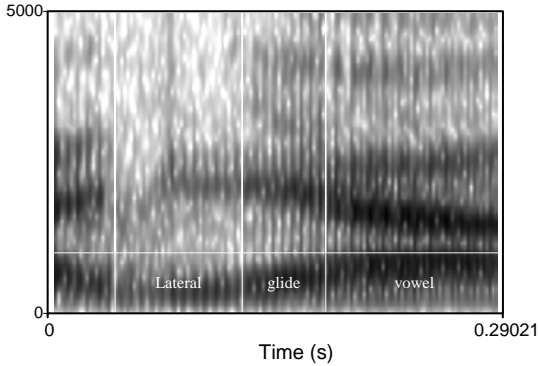


Fig. 1: Spectrogram of the sequence [aʎa], extracted from the word *rallada* “shredded<sub>FEM.SG</sub>” (Bella Vista, female speaker). Vertical lines approximately indicate the lateral, the transition, and vowel

## 4. Results

### 4.1 Palatal lateral

4.1.1 *Quality and duration according to the following vowel and stress.* The first factor that I will consider is the role of the vocalic environment in determining the quality of the palatal lateral, since it has been shown (Bladon & Carbonaro 1978; Quilis et al. 1979) that palatals (as opposed to alveolar laterals) do not significantly vary their quality in different vocalic environments. Table 1 shows that F1 and F2 values for the three vowels tend to overlap, and results of three one-factor ANOVA tests indicate that differences in F1 ( $F(2, 252) = 0.42, p = 0.66$ ), F2 ( $F(2, 252) = 1.38, p = 0.25$ ), and F3 ( $F(2, 252) = 2.93, p = 0.06$ ) are not significant.

Vowel		F1 (Hz)	F2 (Hz)	F3 (Hz)
[a]	<i>N</i>	135	135	135
	Mean	379.4	1816.5	2598.4
	<i>SD</i>	58.4	230.9	347.9
[e]	<i>N</i>	37	37	37
	Mean	327.9	1618.6	2143.7
	<i>SD</i>	116.5	630.9	829.1
[o]	<i>N</i>	80	80	80
	Mean	378.1	1787.9	2516.9
	<i>SD</i>	65.4	237.7	310.1

Table 1: F1-F3 frequencies of the palatal lateral in three vocalic contexts

Formant values may also differ in stressed and unstressed syllables (Quilis et al. 1979) and thus, the role of the variable ‘stress’ was also explored (cf. Table 2). In order to test for possible differences, three paired *t* tests were performed. F1 mean values are significantly higher in unstressed syllables ( $t(269) = 1.98, p < 0.05$ ), while F2, ( $t(269) = 1.38, p > 0.05$ ) and F3 ( $t(269) = 1.11, p > 0.05$ ) were not significant. Higher F1 values in unstressed syllables suggest a more open articulation, while similarities in F2 and F3 under both stress conditions indicate that there are no changes in the location of the constriction.

Stress		F1 (Hz)	F2 (Hz)	F3 (Hz)
Stressed	<i>N</i>	95	95	95
	Mean	369.8	1841.8	2520.2
	<i>SD</i>	58.9	232.6	366.0
Unstressed	<i>N</i>	175	175	175
	Mean	383.1	1798.9	2570.1
	<i>SD</i>	63.2	236.8	335.2

Table 2: *F1-F3 frequencies of the palatal lateral in stressed and unstressed syllables*

If the palatal lateral is being subjected to a change, it may be the case that its duration is affected. Thus, duration of the lateral closure was measured for all speakers, before all vowels, and in stressed and unstressed positions. Since mean values did not differ significantly in the aforementioned contexts, only a summary of results is reported in Table 3.

	Duration (ms)
<i>N</i>	270
Mean	70
Max	90
Min	50
<i>SD</i>	10

Table 3: *Mean duration and summary statistics (all speakers)*

4.1.2 *The role of extralinguistic factors and their relationship with the quality of the palatal lateral.* Table 4 summarizes the results obtained for the first three formants for all the speakers pooled together (see appendix for individual values). An ANOVA test with ‘location’ as a factor indicates that the differences are significant for F1 ( $F(3, 270) = 19.82, p < 0.0001$ ), F2 ( $F(3, 270) = 10.25, p < 0.0001$ ), and F3 ( $F(3, 270) = 8.22, p < 0.0001$ ). Interspeaker differences may be expected from anatomical factors (differences in vocal tract sizes) or from sociolinguistic factors (i.e., women maintaining a variable

associated with prestige, etc.). Thus, two additional ANOVA tests were performed for the group of female speakers and male speakers. Again, interspeaker differences were significant among female (F1:  $F(3, 158) = 29.8, p < 0.0001$ ; F2:  $F(3, 158) = 11.56, p < 0.0001$ ; F3:  $F(3, 158) = 8.63, p < 0.0001$ ) and male speakers (F1:  $F(3, 112) = 6.78, p < 0.0003$ ; F2:  $F(3, 112) = 9.09, p < 0.0001$ ; F3:  $F(3, 112) = 12.53, p < 0.0001$ ).

	F1 (Hz)	F2 (Hz)	F3 (Hz)
<i>N</i>	267	267	267
Mean	377.6	1816.6	2549.1
Max	521.9	2432.5	3383.7
Min	238.7	1293.7	1742.8
<i>SD</i>	60.8	232.2	349.6

Table 4: Mean formant frequencies and summary statistics (all speakers)

Interspeaker variation deserves closer attention, since it may be the case that quality differences in palatal laterals are somehow related to their rate of maintenance. In order to explore this hypothesis, two correlations were calculated: F1 versus rate of maintenance and F2 versus rate of maintenance (see Table 5). The former variables are weakly correlated ( $r = 0.24$ ), whereas the latter are strongly correlated ( $r = 0.65$ ). According to the present results, speakers exhibit a higher F1 (when compared to previous studies), independently of their rate of maintenance, which suggests a more open articulation. On the other hand, frequency of F2 is positively correlated with ‘rate of maintenance’; that is, speakers with a lower rate of maintenance have a lower F2, which suggests a lengthening of the cavity behind the constriction, probably due to a retraction of the tongue body (cf. Narayanan et al. 1997).

Speaker	F1 (Hz)	F2 (Hz)	Rate of maintenance
BAM	351.0	1770.1	100
BAF	354.4	1901.8	100
BVM	376.2	1578.4	84
BVF	461.5	1966.5	100
AM	346.0	1566.1	60
AF	413.8	1615.8	93.1
SMM	428.0	1761.1	84.4
SMF	365.9	1695.9	100

Table 5: Values of F1 and F2 and rate of maintenance for all speakers

#### 4.2. Transition

4.2.1 *Quality and duration according to the following vowel and stress.* In order to test the hypothesis that there is a glide in the CV transition, its quality was analyzed. Table 6 suggests some quality variation according to the following vowel. In order to determine whether those differences are significant, three separate ANOVA tests were performed with ‘vowel’ as a factor group. Differences in the values of F1 ( $F(2, 252) = 10.24$ ;  $p < 0.0001$ ) and F2 ( $F(2, 252) = 5.47$ ,  $p = 0.004$ ) were significant, while F3 was not significantly different ( $F(2, 252) = 2.77$ ,  $p = 0.064$ ).<sup>12</sup>

Vowel		F1 (Hz)	F2 (Hz)	F3 (Hz)
[a]	<i>N</i>	135	135	135
	Mean	455.6	1798.5	2569.5
	<i>SD</i>	70.2	254.8	333.9
[e]	<i>N</i>	37	37	37
	Mean	405.6	1931.4	2694.3
	<i>SD</i>	52.5	264.1	351.2
[o]	<i>N</i>	80	80	80
	Mean	416.4	1759.5	2544.9
	<i>SD</i>	71.2	238.1	302.6

Table 6: *F1-F3 frequencies of transition in three vocalic contexts*

Since it has been shown that vowel formants may vary under different stress conditions (Summers 1987; Martínez Celdrán 1994; de Jong 1995, among others), the role of the variable ‘stress’ was also explored here (cf. Table 7). Results of three paired sample *t* tests indicate that differences were significant for F1 ( $t(269) = 1.21$ ,  $p > 0.05$ ), F2 ( $t(269) = 0.93$ ,  $p > 0.05$ ), and F3 ( $t(269) = 1.41$ ,  $p > 0.05$ ).

Stress		F1 (Hz)	F2 (Hz)	F3 (Hz)
Stressed	<i>N</i>	95	95	95
	Mean	438.1	1827.2	2619.6
	<i>SD</i>	64.8	278.3	366.2
Unstressed	<i>N</i>	175	175	175
	Mean	430.1	1794.1	2557.9
	<i>SD</i>	71.9	247.3	321.3

Table 7: *F1-F3 frequencies of the transition in stressed and unstressed syllables*

<sup>12</sup> Although F3 differences deserve further analyses, Stevens (1998:283) observes that F3 may not play a role in determining vowel quality.

As mentioned, transition duration is a problematic variable, since methods to measure quality and duration of glides vary between researchers. Thus, results obtained can only tentatively be compared with other studies. The mean duration of the transition in Corrientes Spanish is 40 ms (cf. Table 8). As was observed in 4.1.1, mean duration remains constant in different vocalic contexts, and under different stress conditions.

Duration (ms)	
<i>N</i>	270
Mean	40
Max	50
Min	30
<i>SD</i>	10

Table 8: *Mean duration and summary statistics (all speakers)*

4.2.2 *The role of extralinguistic factors and their relationship with the quality of the transition.* Table 9 displays the formant values for all the speakers in the sample. In order to determine whether there were significant differences, ANOVA tests were performed with ‘location’ as a factor. Results indicate that differences were significant for F1 ( $F(3, 270) = 17, p < 0.0001$ ), F2 ( $F(3, 270) = 13.13, p < 0.0001$ ), and F3 ( $F(3, 270) = 4.57, p < 0.01$ ). This holds true also for subgroups of the sample. Differences are significant within the group of female speakers for F1 and F2 ( $F(3, 158) = 14.72, p < 0.0001$ ;  $F(3, 158) = 9.89, p < 0.0001$ , respectively), but not for F3 ( $F(1, 158) = 1.94, p < 0.12$ ). Among the group of male speakers, differences are significant for the first three formants (F1:  $F(1, 112) = 11.7, p < 0.0001$ ; F2:  $F(1, 112) = 15.58, p < 0.0001$ ; F3:  $F(1, 112) = 15.2, p = 0.0001$ ). Again, differences in formant values are expected from possible differences in vocal tract sizes, and in recording conditions. In addition, glide-like transitions are also expected to vary in relation to the vocalic context (see 4.2.1). Thus, if vowel tokens were not equally distributed within the sample (e.g., one group of speakers had more tokens of [ʌa] sequences) formant values would also differ.

	F1 (Hz)	F2 (Hz)	F3 (Hz)
<i>N</i>	260	260	260
Mean	432.2	1805.5	2579.9
Max	588.5	2438.5	3493.4
Min	252.9	1136.7	1807.5
<i>SD</i>	68.9	258.2	338.5

Table 9: *Mean formant frequencies and summary statistics (all speakers)*



### 5. *From palatal laterals to glides*

These data confirm the first hypothesis of the study, that is, that there is a glide in the CV transition. The quality of the element that extends from the release of the closure of the lateral to the steady state of the following vowel is similar to the one reported for the transitions of the palatal lateral in Brazilian Portuguese, and generally falls within the ranges of the glide reported by Borzone de Manrique (1976, 1979), especially for F1 (i.e., 254-471 Hz). However, my values for F2 and F3 are lower. Differences may be attributed to methodological reasons. Borzone de Manrique reports data on [ie] diphthongs, and she acknowledges the fact that formants of the glide are sensitive to vowel environments. Since [e] has a higher F2 and F3 than [a], which turned out to be the most frequent vowel in my sample, a lower F2 and F3 are expected here. Indeed, results in Table 6 indicate that F2 values when the palatal is followed by [e] are 200 Hz higher than in other vocalic contexts. The tendency toward a more centralized articulation observed here is, however, consistent with Borzone de Manrique's (1979:199) observation that "allophones of front vowels are more open in diphthongs than in isolation." Interestingly, transition values for F1 and F2 are close to values reported for glides in onset position in a previous study on Argentine Spanish (Colantoni 2001). Since that study included data about Corrientes Spanish, it would be worth exploring the similarities between glide-like transitions and glides found in onset position, among those speakers who have almost completed the merger.

Transition durations are longer than in Peninsular Spanish (Quilis et al. 1979) and Brazilian Portuguese (Silva 1999). They fall just outside the range (16-38 ms) of the palatal glide in Argentine Spanish (Borzone de Manrique 1979), and are 10 ms shorter than in Peninsular Spanish (Aguilar 1997, 1999). The present results support the hypothesis that transitions eventually become longer and are interpreted as the relevant cue for the segment.

Present findings also confirm the second hypothesis, that the emergence of the glide is not the result of a shortening of the laterals but of a change in its quality (i.e., a more open articulation). The palatal lateral in Corrientes Spanish exhibits a higher F1 than in Peninsular Spanish and in Brazilian Portuguese (Quilis et al. 1979; Silva 1999). Mean values for F2 in Corrientes Spanish are remarkably similar to those reported for Brazilian Portuguese, which are both lower than in Italian (Bladon & Carbonaro 1978) and Peninsular Spanish. A lower F2 suggests a change in the cavity behind the constriction, and possibly that the constriction in Corrientes is taking place further back in the vocal tract. The other characteristics of the palatal lateral coincide with those reported in the literature. Formant patterns do not significantly vary according to the

vocalic environment (Bladon & Carbonaro; Quilis et al.), and the mean duration of the segment coincides with values reported by Quilis et al. for Peninsular Spanish.

Thus, overall results support the general hypotheses that a cue that was already present in the signal survives when the change is completed. Present findings provide evidence for a model (Ohala 1989) that argues that sound change takes place when the listener, during the acquisition process, misinterprets a phonetic cue that is already in the signal. Listeners are, then, interpreting the glide-like element as the relevant feature until it eventually substitutes the palatal lateral.

Finally, an analysis of the role of extralinguistic variables suggests that interspeaker variation in the quality of the lateral cannot be directly correlated with gender effects. The acoustic data confirm previous results (Colantoni 2001) that reveal no significant differences for men and women in the rate of maintenance of the palatal. It is hypothesized then that the emergence of the glide constitutes a change from below (Labov 1994), since there seems to be no awareness in the community, probably due to the similarity between the palatal lateral and the glide. Previous (Colantoni) and present results confirm that 'location' is the relevant factor in predicting the quality of the palatal. Locations to the north of the province show higher F2 values, and higher rates of maintenance. Since those locations coincide with the area where Guarani is more widely spoken, it will be necessary to explore the correlation between the degree of bilingualism and the rate of maintenance of the palatal in future studies.

## 6. *Transitions, coarticulation, and complex segments*

The problem that still needs to be addressed is whether this glide-like transition is an inherent part of the palatal consonant or a result of the CV coarticulation.<sup>13</sup> X-ray evidence (Keating 1988) indicates that palatals involve a longer constriction and two active articulators. Acoustic data suggest that this longer contact surface correlates with a longer duration of the segment (Bladon & Carbonaro 1978). Contrary to Keating, Recasens (1990) and Recasens et al. (1993) demonstrate, also based on articulatory evidence, (a) that alveopalatals (among which palatal laterals and nasals are included) do not have two active articulators, and (b) that they are different from alveolars and true palatals (among which the glide is included). In addition, they explicitly deny the

---

<sup>13</sup> Coarticulation is defined, after Recasens and Pallarès (2001:15), as the gradient approximation of the articulatory characteristics of two adjacent segments, without implying a complete identity of those segments.

presence of a glide. Previous (Silva 1999) and present results indicate the existence of a glide-like CV transition. Indeed, the expected result of reducing the degree of contact in a palatal consonant is a glide (Keating 1988: 83). Moreover, present findings show that the palatal lateral in Corrientes Spanish has a more open, and probably a more retracted, articulation, that is, the consonant is becoming more glide-like (cf. Tables 1 and 6). Thus, we can hypothesize that there is a gradient approximation between the consonant and the following vowel, in other words, that the glide emerges as a result of the coarticulation. Although results from the present study seem to point in the direction of the coarticulation hypothesis, more acoustic and articulatory studies are necessary to confirm it, given the contradictory nature of the current knowledge.

Finally, the interpretation of the glide-like element as belonging either to the palatal or to the transition touches on the question of the phonological representation of Spanish palatals. If the glide is an inherent part of the palatal, a representation such as the one proposed by Lipski (1989) would be adequate. If, on the contrary, we assume that the glide is a separate segment, Carreira's (1988) proposal would account for the data. However, the two-articulator representation is not consistently supported by articulatory data (cf. Recasens 1990; Recasens et al. 1993), and Carreira's representation is based on two problematic assumptions: (a) that the lateral is alveolar (not consistent with articulatory and acoustic evidence); and (b) that the glide is specified in the underlying representation. Previous and present acoustic characterizations do not clearly support the two-segment hypothesis. The lateral differs in quality from the alveolar lateral, and the glide, although present, may be the result of the phonetic implementation, instead of an underlyingly specified segment. Alternative explanations of the phonological patterning of palatals should be explored, such as restricting the sequence of palatals and high vowels by invoking an OCP constraint, and positing an underlying geminate to account for the length of the segment and its special behavior in stress assignment.

## 7. *Conclusions*

Present results confirm the hypothesis that there is a glide in the CV coarticulation in Corrientes Spanish. Values reported indicate that its quality is generally within the range of formant values obtained for the palatal glide in other Spanish dialects, and its duration is slightly longer than in other Spanish dialects.

Palatal laterals in Corrientes Spanish and Peninsular Spanish differ in quality. It may be the case that the palatal lateral in Corrientes Spanish has

always been qualitatively different from the one in Peninsular Spanish, but interspeaker differences seem to indicate that this is not the case. Although all the speakers show a higher F1 than in other Spanish dialects, values for F2 are strongly correlated with the degree of maintenance of the opposition. Thus, we may hypothesize that the observed allophonic pattern is a consequence of both the presence of a glide in the CV transition and a change in quality of the palatal lateral, which mainly involves a more open articulation. As a consequence, the palatal lateral and the glide become increasingly similar. On the other hand, the fact that the transition is longer than in other Spanish dialects contributes to its interpretation as the relevant cue for the segment. Then, the glide emerges as an allophone, and eventually, the glide may strengthen, as in other dialects, and merge with the palatal affricate.

The present conclusions, however, are limited by the conditions in which data collection took place, namely, recording conditions, no strict control of the speech rate, and unbalanced numbers of tokens for each variable measured. In order to strengthen them, further studies are obviously needed. First, it is necessary to enlarge the sample size and complement the acoustic study with articulatory studies on Argentine Spanish speakers. Second, it is important to investigate the acoustic characteristics of palatal laterals in speakers with lower rates of maintenance (less than 50%). Finally, more detailed acoustic descriptions would allow the development of perception studies on the similarities between a palatal lateral and a glide. After that, we will have a better understanding of one of the most extensive sound changes in the *Romania Continua*.

## APPENDIX

Speaker	Statistics	Duration (L)	F1 (L)	F2 (L)	F3 (L)	Duration (T)	F1 (T)	F2 (T)	F3 (T)
BA (m)	<i>n</i>	64	63	63	63	61	59	59	59
	Mean	0.05	351.0	1785.4	2486.9	0.03	372.4	1813.6	2618.8
	<i>SD</i>	0.02	59.9	164.5	301.9	0.01	56.7	165.8	218.9
BA (f)	<i>n</i>	75	74	74	74	74	74	74	74
	Mean	0.07	362.2	1927.7	2461.3	0.04	438.3	1923.4	2659.2
	<i>SD</i>	0.06	47.9	225.0	247.7	0.01	63.4	229.2	316.3
BV (m)	<i>n</i>	20	20	20	20	19	19	19	19
	Mean	0.07	376.2	1596.3	2528.4	0.04	409.9	1694.5	2275.2
	<i>SD</i>	0.02	50.8	125.2	368.6	0.01	39.1	127.9	243.0
BV (f)	<i>n</i>	37	37	37	37	37	37	37	37
	Mean	0.09	448.5	1963.9	2800.1	0.05	509.7	1925.5	2568.1
	<i>SD</i>	0.02	39.34	193.1	270.2	0.07	40.7	256.0	354.6
A (m)	<i>n</i>	13	13	13	13	12	12	12	12
	mean	0.07	346.0	1566.2	2161.9	0.03	400.8	1515.0	2178.8
	<i>SD</i>	0.02	25.2	170.4	582.9	0.01	57.7	168.7	515.9
A (f)	<i>n</i>	26	26	26	26	26	26	26	26
	Mean	0.08	409.5	1664.7	2639.9	0.04	475.6	1527.2	2678.2
	<i>SD</i>	0.03	54.9	288.2	306.9	0.01	40.8	487.2	313.6
SM (m)	<i>n</i>	13	13	13	13	12	12	12	12
	Mean	0.07	417.6	1810.2	2860.9	0.04	476.1	1839.4	2725.9
	<i>SD</i>	0.01	80.8	108.1	386.1	0.01	42.1	155.8	320.6
SM (f)	<i>n</i>	21	21	21	21	21	21	21	21
	Mean	0.07	365.9	1768.1	2558.0	0.04	438.9	1561.0	2499.6
	<i>SD</i>	0.02	55.1	250.1	360.9	0.01	72.9	251.9	366.6

Note: L = lateral; T = transition.

## REFERENCES

- Aguilar, Lourdes. 1997. *De la vocal a la consonante*. Santiago de Compostela: Universidad de Santiago de Compostela.
- , 1999. "Hiatus and Diphthong: Acoustic cues and speech situation differences". *Speech Communication* 28.57-74.
- Bladon, R. A. W & Emanuela Carbonaro. 1978. "Lateral Consonants in Italian". *Journal of Italian Linguistics* 3.43-55.
- Borzone de Manrique, Ana Ma. 1976. "Acoustic Study of /i, u/ in the Spanish Diphthong". *Language and Speech* 19.121-8.
- , 1979. "Acoustic Analysis of the Spanish Diphthongs". *Phonetica* 36.194-206.
- Carreira, Maria. 1988. "The Structure of Palatal Consonants in Spanish". *Chicago Linguistic Society* 24.73-87.

- Colantoni, Laura. 2001. *Mergers, Chain Shifts and Dissimilatory Processes: Palatals and rhotics in Argentine Spanish*. Ph.D. dissertation, University of Minnesota.
- de Jong, Kenneth J. 1995. "The Supraglottal Articulation of Prominence in English: Linguistic stress as localized hyperarticulation". *Journal of the Acoustical Society of America* 97.491-504.
- De los Heros Diez Canseco, Susana. 1997. *Language Variation: The influence of speakers' attitudes and gender on sociolinguistic variables in the Spanish of Cusco, Peru*. Ph.D. dissertation, University of Pittsburgh.
- Fant, Gunnar. 1960. *Acoustic Theory of Speech Production*. The Hague: Mouton.
- , 1968. "Analysis and Synthesis of Speech Processes". *Manual of Phonetics* ed. by Bertil Malmberg, 173-277. Amsterdam: North-Holland.
- Keating, Patricia. 1988. "Palatals as Complex Segments: X-ray evidence". *UCLA Working Papers in Phonetics* 69.77-91.
- Kovacci, Ofelia. 1987. *Documentos del PREDAL Argentina 2. Atlas lingüístico-antropológico de la República Argentina*. Buenos Aires: Ministerio de Educación y Justicia.
- Labov, William. 1994. *Principles of Linguistic Change: Internal factors*. Oxford: Blackwell.
- Ladefoged, Peter & Ian Maddieson. 1996. *The Sounds of the World's Languages*. Oxford: Blackwell.
- Lipski, John. 1989. "Spanish Yeísmo and the Palatal Resonants: Towards a unified analysis". *Probus* 1.211-223.
- , 1994. *Latin American Spanish*. New York: Longman.
- Lloyd, Paul. 1993. *Del latín al español*. Madrid: Gredos.
- Maddieson, Ian. 1984. *Patterns of Sound*. Cambridge: Cambridge University Press.
- Martínez Celdrán, Eugenio. 1994. *Fonética*. Barcelona: Teide.
- McCarthy, John. 1986. "OCP Effects: Gemination and antigemination". *Linguistic Inquiry* 17.207-63.
- Microsoft Excel. Version 10.0.2614. Microsoft Corporation.
- Narayanan, Shrikanth, Abeer Alwan, & Katherine Haker. 1997. "Towards Articulatory-Acoustic Models for Liquid Approximants Based on MRI and EPG Data. Part I. The laterals". *Journal of the Acoustical Society of America* 101.1064-1077.
- Navarro Tomás, Tomás. 1970. *Manual de pronunciación española*. Madrid: CSIC.
- Nyrop, Kristoffer R. 1923. *Manuel phonétique du français parlé*. Copenhagen: Copenhagen & Kristiania.
- Ohala, John. 1989. "Sound Change Is Drawn From a Pool of Synchronic Variation". *Language Change: Contribution to the study of its causes* ed. by Leiv Breivik & Jahr Ernst, 173-198. Berlin: de Gruyter.
- Penny, Ralph. 2000. *Variation and Change in Spanish*. Cambridge: Cambridge University Press.
- Praat: Doing phonetics by computer. Version 4.0.41. Paul Boersma & David Weenink, Amsterdam: University of Amsterdam. (Available online: <http://www.praat.org>)
- Quilis, Antonio. 1993. *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- Quilis, Antonio, M. Esgueva, L. Gutiérrez Araus, & M. Cantarero. 1979. "Características acústicas de las consonantes laterales españolas". *Lingüística Española Actual* 1.233-343.
- Recasens, Daniel. 1990. "The Articulatory Characteristics of Palatal Consonants". *Journal of Phonetics* 18.267-80.
- Recasens, Daniel, Edda Farnetani, Jordi Fontdevila, & M. Dolores Pallarès. 1993. "An Electropalatographic Study of Alveolar and Palatal Consonants in Catalan and Italian". *Language and Speech* 36.213-234.

- Recasens, Daniel & M. Dolores Pallarès. 2001. *De la fonètica a la fonologia*. Barcelona: Ariel.
- SAS. Version 8.2. SAS Institute, Cary, NC.
- Silva, Adelaide. 1999. "Caracterização acústica de [R], [r], [L] e [ʎ] nos dados de um informante paulistano". *Cadernos de Estudos Linguísticos* 37:51-68.
- Stevens, Kenneth. 1998. *Acoustic Phonetics*. Cambridge, MA: MIT Press.
- Summers, W. Van. 1987. "Effects of Stress and Final-Consonant Voicing on Vowel Production: Articulatory and acoustic analyses". *Journal of the Acoustical Society of America* 82:847-63.
- Vidal de Battini, Berta. 1964. *El español de la Argentina*. Buenos Aires: Consejo Nacional de Educación.
- Whitley, Stanley. 1995. "Spanish Glides, Hiatus and Conjunction Lowering". *Hispanic Linguistics* 6-7:355-85.

# INTERVOCALIC VELAR NASALS IN GALICIAN\*

SONIA COLINA  
*Arizona State University*

## 0. *Introduction*

The syllabic affiliation of intervocalic velar nasals in Galician, for example, *unha* [úŋa] “a, one<sub>FEM</sub>” has been a controversial topic among scholars for at least three decades. All solutions proposed to date (onset or coda affiliation) are costly from a phonological point of view as they run counter to well-attested principles of phonological theory, such as Structure Preservation (Kaisse & Shaw 1985; Kiparsky 1985; Mohanan 1986) and syllabic-markedness generalizations (\*VC.V). In this chapter, I argue that an analysis based on an underlying velar nasal does not encounter any of the difficulties of previous accounts. Consideration of historical and cross-linguistic facts as part of an OT account offers additional support for the underlying nature of the velar and sheds light on aspects of the grammars of related languages like Portuguese and Gascon. The chapter is organized as follows: after the presentation of the data (section 1) and a brief summary of previous accounts (section 2), the proposal and its advantages are introduced in general terms in 3.1; in 3.2 a detailed optimality-theoretic analysis of diachronic and synchronic facts is presented. The underlying velars are motivated also in that section. Section 4 concludes the chapter.

## 1. *Intervocalic velar nasals in Galician*

In Galician,<sup>1</sup> a three-way phonemic distinction in nasals (bilabial, alveolar, and palatal) as in (1) is neutralized in the rhyme (2). Rhyme nasals surface with a velar point of articulation (the phenomenon is similar to that described in some varieties of Spanish, e.g., Cuban, Lipski 1994).

---

\* I would like to thank Sonya Bird and Rachel Hayes for helpful discussion on gemination and ambisyllabicity. My gratitude also goes to José Ignacio Hualde and Fernando Martínez-Gil for useful comments and discussion and to three anonymous reviewers. All remaining errors are my own.

<sup>1</sup> Galician is a Romance language of northwestern Spain, closely related to Portuguese and spoken by approximately two and a half million people.



(1) Martínez-Gil (1993) (Moaña dialect):

<i>mora</i>	[mɔ.ra]	“blackberry”
<i>nora</i>	[nɔ.ra]	“daughter-in-law”
<i>ñopa</i>	[ɲɔ.pa]	“short-sighted <sub>FEM</sub> ”
<i>cama</i>	[ká.ma]	“bed”
<i>cana</i>	[ká.na]	“cane”
<i>caña</i>	[ká.ɲa]	“type of brandy”

The simplest examples of velarization can be seen in word-final position before a pause as in (2). In preconsonantal position, there is often co-articulation (e.g., [kaɲn saɲ] “healthy dog”).

(2) <i>lan</i>	[laɲ]	“wool”	*[lam], [lan], [laɲ]
<i>son</i>	[soɲ]	“sound”	*[som], [son], [soɲ]
<i>ben</i>	[beɲ]	“well”	
<i>xoven</i>	[soβiɲ]	“young”	
<i>irmán</i>	[irmáɲ]	“brother”	

Rhyme nasals surface as velars prevocally across compounds and across words. The same happens before word-final epenthetic /e/,<sup>2</sup> as in (3):

(3) <i>benestar</i>	/benestár/	[be.ɲis.tár]	*[be.nis.tár]	“well-being”
<i>tren alemán</i>	/trenalemán/	[tre.ɲa.li.máɲ]		“German train”
<i>alemán</i>	/alemán/	[a.li.má.ɲi]	*[a.li.má.ni]	“German <sub>MASC</sub> ”

Prefix-final, prevocalic nasals are usually not velarized, as seen in *in+útil* [inútil] \*[iɲú.til] “useless”; exceptional data such as *inhumano* /in + uman + o/ [i.ɲu.má.nu] “inhumane” can be explained through the influence of the spelling *nh*.

In sum, in lexical-phonological terms, nasal velarization is a lexical rule that applies before compounding as well as postlexically. Otherwise, nasals are not velarized before suffixes, such as gender or word markers, as in (4):

(4) a. <i>cana</i>	/kan + a/	[ká.na]	~*[ká.ɲa]	“cane”
<i>cano</i>	/kan + o/	[ká.nu]	~*[ká.ɲa]	“pipe, gutter”
<i>pano</i>	/pan + o/	[pá.nu]	~*[pá.ɲu]	“cloth”
b. <i>lambón</i>	/lamb + on/	[lambóɲ]		“sweet-toothed <sub>MASC</sub> ”
<i>lambona</i>	/lamb + on + a/	[lambóna]		“sweet-toothed <sub>FEM</sub> ”
<i>alemán</i>	/aleman/	[alimáɲ]		“German <sub>MASC</sub> ”
<i>alemana</i>	/aleman + a/	[alimána]		“German <sub>FEM</sub> ”

<sup>2</sup> Word-final epenthesis is an optional process “introduced to satisfy foot binarity at the postlexical level of the intonational phrase” (Martínez-Gil 1997:328).

An important group of words, however, appear to undergo velarization before suffixation. These consist of the feminine forms of the indefinite article and its derivatives, as seen in (5) and (6).

(5) <i>un</i>	/un/	[uŋ]	“a, one <sub>MASC</sub> ”
<i>algún</i>	/alx + un/	[al.xúŋ]	“some, any <sub>MASC</sub> ”
<i>nengún</i>	/neng + un/	[niŋ.gúŋ]	“none <sub>MASC</sub> ”
<i>dun</i>	/de un/	[duŋ]	“of-a <sub>MASC</sub> ”
<i>cun</i>	/kon un/	[kuŋ]	“with-a <sub>MASC</sub> ”
<i>nun</i>	/en un/	[nuŋ]	“in-a <sub>MASC</sub> ”
(6) <i>unha</i>	/un + a/	[úŋa] <sup>3</sup>	“a, one <sub>FEM</sub> ”
<i>algunha</i>	/alx + un + a/	[al.xúŋa]	“a, one <sub>FEM</sub> ”
<i>nengunha</i>	/neng + un + a/	[niŋ.gúŋa]	“none <sub>FEM</sub> ”
<i>dunha</i>	/de un + a/	[dúŋa]	“of-a <sub>FEM</sub> ”
<i>cunha</i>	/kon un + a/	[kúŋa]	“with-a <sub>FEM</sub> ”
<i>nunha</i>	/en un + a/	[núŋa]	“with-a <sub>FEM</sub> ”

## 2. *Previous accounts*

While most scholars seem to agree that any valid account of the forms in (6) must refer to some type of exceptionality (but see Castro 1989), there is in general no agreement in the literature as to what the syllabic affiliation of the velar nasal in these examples might be. Two positions have been adopted in the past to explain the syllabification facts: [ŋ] is in the onset and [ŋ] is in the coda. The issue bears on apparent violations of or exceptions to well-attested generalizations and/or universals of phonological theory. Those who argue that the velar nasal is in the coda still need to explain the resulting onsetless syllable and the extremely marked syllabification pattern in which an intervocalic consonant is syllabified as the coda of the first syllable, (e.g., [uŋ.a]). Thus, while the coda analysis captures the generalization that in Galician rhyme nasals are velars, it must also contend with an extremely rare, mostly unattested pattern of syllabification. The onset proposal does not incur syllabic markedness costs; however, the difficulties encountered are not any less serious. Because according to these proposals, Galician does not have underlying velar nasals and nasal velarization is a lexical rule (cf. Martínez-Gil 1993), the presence of a nonunderlying segment in the lexical domain constitutes a violation of Structure Preservation (Kaisse & Shaw 1985;

<sup>3</sup> Syllabification of the last two syllables in (6) is not indicated in the forms where syllabic affiliation is in question.

Kiparksy 1985; Mohanan 1986).<sup>4</sup> In other words, under the onset analysis, the Galician data in (6) appear to contradict a well-known generalization of phonological theory—that rules with a lexical application do not normally introduce segments that are not part of the phonemic inventory of the language. A third possibility, to be adopted in part here, is to allow velar nasals to be part of the phonemic inventory of the language, thus arguing on the basis of these examples that velar nasals are phonemic in Galician, despite the existence of only one known contrast [uŋ.a] “a, one<sub>FEM</sub>” ~ [u.na] “join<sub>3SG.PRES.SUBJ.</sub>”

There is extensive descriptive work by Galician scholars on the onset/coda debate. In most cases, however, this type of work is purely descriptive, presenting no arguments for the particular position adopted (e.g., Porto Dapena 1976 and Carballo Calero 1979 for coda; Veiga 1976 and Alvarez, Regueira, & Monteagudo 1986 for onset; see also Dubert 1998 for a comprehensive list of references on the descriptive literature). Despite the lack of specific phonological argumentation, these studies are of value because they suggest the inability of native speakers (most of this literature is written by Galician scholars who are also native speakers of the language) to decide on the syllabic affiliation of velar nasals. An informal consultation of several native speakers about their syllabification intuitions confirms this view and supports an analysis in which the nasals in (6) occupy two syllabic positions (cf. section 3.2).

Within a lexical phonology framework, and also arguing for onset affiliation, Castro (1989) explains that the difference between [una] and [uŋa] is the result of morphological adjunction taking place in different lexical strata. She claims that in [una] the thematic vowel attaches to the verb root in the lexical stratum (level 1), leaving the nasal in onset position before velarization applies; in [uŋa], however, the gender marker *-a* is attached postlexically, after nasal velarization. In addition to the difficulty of motivating word-marker adjunction postlexically (in most languages it seems to be a lexical process), Castro's proposal cannot satisfactorily account for forms like *lambón* and *lambona* in (4b), which should also have a velar nasal. She argues that forms with an alveolar nasal before the feminine gender marker have an empty V in the masculine (e.g., *lambónV*). As Dubert (1998) also points out, Castro's solution is ad hoc, since there is no evidence to support the presence of V in *lambón* (and not in *un*). Furthermore, under this account, the more numerous

---

<sup>4</sup> Martínez-Gil (1993) argues that Structure Preservation must be modified to allow for the Galician data and a few other cases of apparent violations. Under the current proposal, this is not necessary, at least for the Galician data.

forms—those with alveolar nasals before the suffix—are treated as exceptional, whereas *un/unha* and its derivatives are unexceptional.

Lipski (1976) reviews various phonological, historical, and cross-linguistic factors that may play a role in accounting for intervocalic velar nasals in Galician. Although he does not take a stance on the onset/coda debate, he makes an important contribution by pointing out that the solution to the problem lies in “the interaction of several factors” (1976:191). I argue that this is exactly the case and that for this reason an optimality-theoretic account based on constraint interaction, like the one proposed here, is well suited to account for the syllabic affiliation of velar nasals in Galician.

Dubert (1998) presents a somewhat different proposal. In support of the onset analysis, he argues for an underlying velar nasal for the relevant forms (/uŋ/, /u.ŋa/) which is then resyllabified after suffixation. This analysis faces several difficulties among which is the loss of an important generalization—rhyme nasals are velar in Galician—and the extremely marked status of an onset velar nasal. Dubert’s proposal, however, captures the need to mark these forms as exceptional by positing underlying velar nasals in the relevant cases. I will argue that it is possible to retain the advantages of Dubert’s analysis, while also satisfying syllabic-structure conditions and generalizations.

### 3. *Analysis*

#### 3.1 *Basic proposal: Underlying velar nasals*

The analysis I propose in this chapter argues for the exceptionality of *unha* forms (vs. *alemana*). I claim that the forms in (6) differ from those in (4b) in that the relevant nasals in (6) are underlyingly velar. In addition, the current proposal involves velars with a multiply linked supralaryngeal node, more specifically, a surface geminate. The geminate is the result of the assimilation of an epenthetic onset to the point of articulation of the preceding nasal.

This analysis does not suffer from any of the drawbacks of previous accounts. The highly marked VC.V syllabification pattern and the onsetless syllable of the coda analysis are avoided because the second half of the geminate serves as an onset for the second syllable, [uŋ.ŋa]. At the same time, the output geminate serves to preserve the generalization that coda nasals are velar. In addition, in an analysis in which velar nasals are phonemic, there is no violation of Structure Preservation, given that velar nasals are part of the phonological inventory of the language. The proposed account also captures native-speaker intuitions that the forms in (6) are to some extent exceptional (expressed through the underlying point of articulation). Such exceptionality, however, is not random, as underlying velars appear in forms that lost

intervocalic nasals in their development from Latin. Finally, the current proposal accounts for ambiguous syllabification and native speakers' difficulty in assigning velar nasals to either coda or onset. Another advantage of the analysis proposed is that, as will be shown in section 3.2, it sheds light on related developments in other Romance languages.

Nonetheless, exceptionality and gemination alone are not sufficient to account for velar nasals in Galician. Given that any successful account of velar nasals must be able to capture various interacting factors, OT (through constraint interaction and ranking) presents itself as the right type of framework for formalizing these facts. As I will show, in an OT framework, underlying velars are not established a priori, but they are the result of the ranking of the constraints (in conjunction with Lexicon Optimization, as part of a theory of acquisition).

### *3.2 Historical developments and cross-linguistic evidence*

In what follows I show that consideration of historical and cross-linguistic facts, analyzed within an OT framework, provides further evidence for the current proposal. Since OT is an inherently typological model of language competence, the analysis and claims made about the grammar of Galician with regard to intervocalic velar nasals inevitably incorporate claims about the grammars of other languages, in particularly of those where similar processes are active. Consequently, the ease with which a particular analysis can account for related facts in other languages is a decisive factor in evaluating its adequacy. I show that the OT analysis proposed for the Galician data can account for diachronic and synchronic data in Portuguese and Gascon, shedding light on the grammars of those languages.

Although the main concern of this chapter is not to explain diachronic data, a few words must be said about optimality-theoretic accounts of language change, especially since the analysis of the data presented here is couched within that framework. Within an OT framework, I espouse a view of historical change as constraint reranking, which results from intermediate stages characterized by variable ranking ("diachronic change via synchronic variation" (McCarthy 2002:229)); in other words, different synchronic stages of the history of a language are the result of different rankings of the constraints; a difference in ranking is the final stage arising as a consequence of a period of unstable, variable rankings of the relevant constraints. Thus the proposal presented in this section looks at several diachronic stages as snapshots in the evolution of the language, without further inquiry into how the

language proceeded from one stage to another. I consider that to be beyond the goal of this chapter.

Another issue that needs to be addressed with regard to an OT analysis of intervocalic velar nasals in Galician is that of their underlying representation. As I discuss in more detail below, OT places no restrictions on the form of the input. Underlying representations are a natural consequence of the constraints and constraint ranking (cf. Richness of the Base, McCarthy 2002) in conjunction with Lexicon Optimization, a learning strategy that, in the absence of alternations, helps the learner select the underlying forms that most harmonically map onto the output. What I refer to here in more traditional terms as ‘underlying velar nasals’ is in fact an epiphenomenon of constraint ranking.

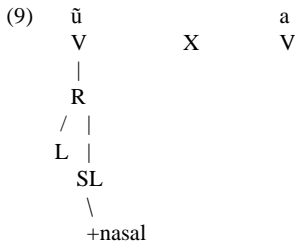
The forms exhibiting ambiguous syllabification of velar nasals in Galician happen to be those that resulted from the loss of Latin intervocalic nasals among Galician and Portuguese dialects and other Romance-speaking areas (Meyer-Lübke 1921). In Galician, loss of intervocalic nasals initially resulted in vowel hiatus, with nasalization of the preceding vowel; the nasal later reappeared as a velar ([uŋa]). In Portuguese and Gascon, Latin’s intervocalic nasals were also lost, leaving behind a nasalized vowel (Williams 1962; Rohlf’s 1970); standard Portuguese forms like [ũa] became [uma], whereas in Gascon they remained as [ya].

(7)	Latin	<u>Galician/Portuguese/Gascon</u> (intermediate stage)	<u>Galician</u> <u>Portuguese</u> <u>Gascon</u>		
	unam	ũa	uŋa	uma	ya

I assume the feature-geometrical representation in (8) (Clements 1985; Sagey 1986).

(8)	ũ	n	a
	V	X	V
	R	R	
	/	\	
	L	L	
	SL	SL	
	\	/	
	+nasal	P A	
		+coronal	

I claim that the deletion of intervocalic nasals in Latin left behind an empty consonantal slot, with nasality preserved on the preceding vowel, as in (9):



In OT terms, the deletion of intervocalic nasals is the result of a conflict between ease of articulation, facilitated by deletion of the point of articulation of the nasal and preservation of nasality on the vowel, and faithfulness to the underlying representation. Ease-of-articulation effects can be obtained through LAZY (Kirchner 1998). LAZY is a family of constraints regulating minimization of effort. The LAZY constraint associated with segments that require greater effort is ranked higher than the one associated with segments requiring less effort. In the Galician case, LAZY VnV (the effort required to pronounce intervocalic nasals) dominates LAZY VxV (the effort required for an intervocalic, placeless slot). For the sake of brevity, here I use ‘LAZY’ to refer to the relevant LAZY constraints. Alternatively, the same effects can be argued to result from an AGREE (Closure) constraint that requires adjacent segments to have the same degree of closure—an intervocalic consonant would violate this constraint because of the obstruction to the airflow necessary for the articulation of a consonant. An intervocalic placeless slot does not incur a violation of AGREE.

Additional relevant constraints are in (10):

- (10) MAX-IO: Input segments must have output correspondents (no deletion).  
 MAX-IO (Place): The place node of an input segments must have a place node correspondent in the output (no deletion of place features).  
 IDENT-IO (Nasal): The value of the feature [nasal] of an input segment must be preserved in its output correspondent.  
 Linearity-IO: The output reflects the precedence structure of the input, and vice versa (Pater 1999).

LAZY and MAX-IO are top-ranked constraints, since neither total faithfulness to the input nor deletion of the entire segment are possible options. The unspecified point of articulation in the empty slot, as in (9), indicates that

MAX-IO (Place) must be ranked lower than MAX-IO, since an empty skeletal slot (no PA node) is preferred to total deletion; preservation of nasality on the preceding vowel demonstrates that IDENT-IO (Nasal) dominates Linearity-IO, since alteration of the precedence structure of the input (nasality on the preceding vowel rather than on the empty slot) is preferable to no nasality at all. Thus, at one point in the history of Galician, Portuguese, and Gascon, the constraint ranking in (11) was operative:

(11) LAZY, MAX-IO >> MAX-IO (Place), IDENT-IO (Nasal) >> Linearity-IO

(12)/unam/ [ũ\_a]

	LAZY	MAX-IO	MAX-IO(Place)	IDENT-IO (Nasal)	Linearity-IO
a. unam	*!				
b. ua		*!	*		
c. <i>ũ_a</i>			*		*
d. u_a			*	*!	

As the tableau in (12) illustrates, candidate (c) is the optimal candidate and the output, because (a) and (b) violate the top-ranked LAZY and MAX-IO; (d), without nasalization on the vowel, incurs an additional violation of IDENT-IO (Nasal), which dominates Linearity-IO (violated by (c)).<sup>5</sup> It is reasonable to assume that the ranking in (12) would eventually lead to the postulation of /*ũ\_a*/ as the underlying representation (i.e., phonemicization of [ũ\_a]), as this would be the most harmonic mapping for the output under Lexicon Optimization. /*ũ\_a*/ is in turn the initial input to the ranking variations that would produce the Galician and Portuguese forms in (15) and (19).

In later diachronic stages, Galician and Portuguese eventually recovered the point of articulation of the intervocalic nasal. In Galician the back quality of [u] produced a velar, [+back, +hi], whereas in Portuguese, [+round] produced a bilabial nasal. While Gascon, in particular Aranese Gascon, allows empty slots (cf. later discussion and Hualde 1992 for evidence in favor of the presence of empty consonantal slots in Gascon), Galician and Portuguese do not, which caused the reappearance of the nasal. As seen in (13a), in OT this is the consequence of the emergence of a new ranking in Galician and Portuguese in which a constraint against empty slots or placeless consonants (\*X)

<sup>5</sup> Some authors have suggested that deletion of Latin intervocalic nasals in strong onset position is facilitated by ambisyllabification, which would place part of the nasal in a weak coda position (Hajek 1997:200); if this were to be the case, NO CODA would replace LAZY. Under this hypothesis, languages with no deletion failed to undergo ambisyllabification.



dominates LAZY (vs. LAZY >> \*X, the ranking that produced the loss of intervocalic nasals in Latin). Gascon retained LAZY >> \*X (13b).

(13) a. Galician, Portuguese: \*X >> LAZY

Galician, Portuguese: \*X >> MAX-IO (Place), IDENT-IO (Nasal) >> Linearity-IO

b. Gascon: LAZY >> \*X

Gascon: IDENT-IO (Place), MAX-IO (Nasal) >> Linearity-IO >> \*X

It is necessary now to explain why Galician chooses a velar, while Portuguese prefers a bilabial nasal. Note that Portuguese does not have any coda nasals, nor does it allow the presence of velar nasals in the onset. This leaves no choice but to spread the [+labial]/[+round] feature of /u/. In OT terms, a constraint against velar nasals (\*ŋ) is highly ranked in Portuguese. In (15) and (19) I propose /ũ\_a/ (/ỹ\_a / in (26)) as the input because this is the most harmonic mapping for the output (given the set of constraints and constraint rankings) at the time the change took place. Any other potential input would produce the same output under the established ranking.

(14) Portuguese: \*X, \*ŋ >> MAX-IO (Place), IDENT-IO (Nasal) >> Linearity-IO

(15) Portuguese: /ũ\_a/[ũma]

	*X	*ŋ	MAX-IO (Place)	IDENT-IO (Nasal)	Linearity-IO
a. ũma					**
b. ũ_a	*!				
c. ũ.ŋa		*!			**
d. ũŋ.a		*!			**

In (15) candidate (b) violates \*X because of the empty slot; since (c) and (d) contain violations of the highly ranked constraint \*ŋ, (a) is the winner. It is important to note that the spreading of the nasal and back/round features of [u] to the empty slot constitute Linearity-IO violations, not IDENT-IO, MAX-IO, or DEP-IO because no feature is altered, deleted, or inserted (it is only its precedence structure that is modified).<sup>6</sup> Additional support for this account can be found in forms in which the intervocalic nasal is preceded by a high, front vowel, for example, Latin VINUM > vi\_o > Portuguese *vinho* [bino] (Williams 1962). As expected, the [+high] and [-round] features of the vowel result in a palatal nasal through a violation of Linearity.

<sup>6</sup> Also note that the output vowel could be nasalized or not. Many languages show variation in this respect. This has no effect on the results of the analysis.

In Galician, syllabification is a necessary consideration in candidate evaluation. [ŋ] is not permitted in the onset, although it is well formed as a coda. A form like [ũŋ.a], however, would violate the highly-ranked ONSET constraint. I propose that the conflict is resolved by means of the ranking  $*_{\text{Ons}}[\eta]$ , ONSET  $\gg$  DEP-IO,  $*\eta$ , which results in consonant epenthesis. The epenthetic consonant takes its features from the velar nasal and thus the resulting double-linked SL node satisfies the constraint against velar nasals in the onset:  $*\text{ALIGN-LEFT}(\eta, \sigma)$  (abbreviated as  $*_{\text{Ons}}[\eta]$ ).  $*_{\text{Ons}}[\eta]$  is defined, in accordance with the crisp version of alignment (McCarthy & Prince 1993:10), as a constraint against a velar nasal being aligned with the syllable onset. Since doubly linked [ŋ] is not ‘crisply’ aligned with the onset, it does not violate  $*_{\text{Ons}}[\eta]$ .<sup>7</sup>

(16) Galician:  $*X, *_{\text{Ons}}[\eta], \text{ONSET} \gg \text{DEP-IO(C)}, \text{Linearity-IO}$

The question remaining, however, is why does Galician not select [ũma] in order to avoid either an onsetless syllable or an onset velar, without having to resort to epenthesis? The answer can be found by looking at the rest of the phonology. As mentioned before, rhyme nasals are velarized in Galician, especially in word-final position (this obviously does not happen in Portuguese, which does not have coda nasals). Therefore /un/, the masculine singular form of the indefinite article, is realized as [uŋ]. The existence of this form in isolation (output) puts analogical pressure on the feminine to resemble the output of the masculine base as much as possible and thus, a feminine with the same point of articulation as the masculine is preferred. In OT terms, the output-to-output constraint IDENT-BA in (17), requiring identity of the output of a derived form to the output of the base in isolation, is highly ranked and it dominates both DEP-IO and Linearity-IO (18), resulting in the selection of [uŋ.ŋa], as seen in (19). This account also captures the observation made by some scholars that the masculine form of the article must play a role in this phenomenon (Lipski 1976; Porto Dapena 1976).

---

<sup>7</sup> An alternative formalization of these facts consists of replacing the negative/markedness alignment constraint  $*_{\text{Ons}}[\eta]$  with a positive, plain alignment constraint, ALIGN-RIGHT ( $\eta, \sigma$ ), which tolerates noncrisp edges (see Ito & Mester 1994 for definitions of plain alignment and noncrisp edges) and requires velar nasals to be in the coda. Doubly linked segments are the best way to satisfy ALIGN-RIGHT ( $\eta, \sigma$ ), and high-ranked ONSET (= ALIGN-L (C,  $\sigma$ )).

- (17) IDENT-BA[ $u\eta$ ]: Given an input structure [X Y] output candidates are evaluated for how well they match [X] and [Y] if the latter occur as independent words (Benua 1995; Kenstowicz 1996).

ONSET: All syllables have onsets.

- (18) Galician: ONSET, \*X, \* $_{\text{Ons}}[\eta]$ , IDENT-BA[ $u\eta$ ] >> DEP-IO(C), Linearity-IO

- (19) Galician: / $\tilde{u}_a/$  / $[\tilde{u}\eta. \eta a]$

	ONSET	*X	* $_{\text{Ons}}[\eta]$	IDENT-BA [ $u\eta$ ]	DEP-IO(C)	Linearity-IO
a. $\tilde{u}.ma$				*!		**
b. $\tilde{u}._a$		*!		*		
c. $\tilde{u}\eta.a$	*!					**
d. $\tilde{u}.\eta a$			*!			**
e. $\tilde{u}\eta.\eta a$					*	**

Candidates (b), (c), and (d) in (19) all violate one of the top-ranked constraints (\*X, ONSET and \* $_{\text{Ons}}[\eta]$ , respectively). (a) is also eliminated because it incurs a violation of IDENT-BA [ $u\eta$ ] (not incurred by the winner). Note that the winner, candidate (e), does not violate \* $_{\text{Ons}}[\eta]$  because the velar nasal, occupying two syllabic positions, is not aligned with the onset. It incurs a DEP-IO (C) violation, but this constraint is ranked lower than the ones violated by (a)-(d). Lack of alternation eventually leads to phonemicization of the velar point of articulation in / $\tilde{u}\eta a$ /. In OT terms, / $u\eta a$ / is the input that most harmonically maps onto the output and thus is posited by the Galician learner as the underlying representation. Any other possible input subjected to the constraints and constraint ranking would produce the same output.<sup>8</sup> Preliminary phonetic evidence also appears to have been found for different types of velar nasals in Galician (underlying in *unha* and derived in *alemani*) and for the presence of a geminate on the surface, as the underlying velars are significantly longer than velars in other positions (Colina & Díaz-Campos 2003). Furthermore, Colina and Díaz-Campos also report that in some instances a glottal stop surfaces in place of the second half of the geminate. This lends support to the consonant-epenthesis proposal as epenthetic consonants are often the least marked consonants (glottal stops, aspiration, segments without a supralaryngeal node), reflecting the emergence of the unmarked when faithfulness to underlyingly specified features is not at stake.

<sup>8</sup> Galician has some exceptional forms in which the intervocalic nasal was not recovered, for example, Latin LUNAM > *lu\_a* > Galician *lua* ‘moon.’ Although Standard Portuguese shares the same development (Portuguese *lua*, \**luma*), *luma* also exists in some dialects (Williams 1962:83).

Consonant epenthesis, however, is not always resorted to in Galician to create a syllable onset. The constraints and constraint ranking relevant to the specific cases account for lack of consonant epenthesis. For instance, in word-initial position, as in *irmán* “brother,” the ranking ALIGN-L >> ONSET eliminates an output with C-epenthesis \**Cirmán*, as that would violate ALIGN-L, a constraint that requires left alignment of the morphological word with the prosodic word. In sequences of VCV, as for instance in *cana* in (1), the high rankings of both NO CODA and ONSET conspire to select V.CV, thus obviating the need to violate DEP-IO(C) through consonantal epenthesis. Morpheme internally, VV is resolved as either a diphthong or a hiatus; in the latter case, epenthesis is avoided because the need to preserve morpheme integrity (CONTIGUITY) is more important than to have a syllable with an onset (CONTIGUITY >> ONSET).

An important advantage of the current analysis is that it accounts for the presence of an alveolar nasal in [alimá.na]. The crucial difference between [alimá.na] and [uŋ.ŋa] rests in the absence of the empty slot in the former, since /alemán/ [alimán] ~ /alimána/ [alimá.na] does not result from deletion of intervocalic nasals (Latin ALEMANE > *aleman*). Thus, there is no possible violation of \*X. The reader will recall that a velar nasal is chosen to fill the empty slot and that a velar point of articulation is preferred to a labial one, because of the presence of a velar in the masculine form and of the constraint requiring identity between base and derived form. In [alema.na], however, \*X is vacuously satisfied (no deletion of intervocalic nasals), so only velarization constraints are relevant. It is not the objective of this chapter to provide an account of velarization in Galician, which is a rather complex process in its own right given its interaction with the morphology and syntax. I will, however, sketch the basic lines of a possible analysis insofar as it is necessary to explain the difference between [alimá.na] and [uŋ.ŋa]. For the masculine, markedness constraints for the coda (a velar nasal is less marked than an alveolar one (\*<sub>Coda</sub>[n] >> \*<sub>Coda</sub>[ŋ]) dominate faithfulness ((\*<sub>Coda</sub>[n] >> \*<sub>Coda</sub>[ŋ], IDENT-IO(Place)), as seen in (20).

(20) /alemán/ [alimán]

	* <sub>Coda</sub> [n]	IDENT-IO(Place)	* <sub>Coda</sub> [ŋ]
a. alimán	*!		
☞ b. alimánŋ		*	*

For the feminine in (21), however, since the nasal is in the onset position, faithfulness to the underlying representation (IDENT-IO(Place)) selects

[alimána]. The losing candidate also incurs a violation of the highly ranked  $*_{\text{Ons}}[\eta]$ .

(21)/alemána/ [alimána]

	$*_{\text{Coda}}[n]$	IDENT-IO(Place)	$*_{\text{Coda}}[\eta]$
a. alimána			
b. alimána		*!	

The same ranking explains the ill-formedness of *unha*  $*[una]$ , as seen in (22).

(22)/uŋa/ [uŋ.ŋa]

	$*_{\text{Coda}}[n]$	IDENT-IO(Place)	$*_{\text{Coda}}[\eta]$
*a. u.na		*!	
b. u.ŋa			

In this case, in addition,  $*_{\text{Ons}}[\eta]$  becomes relevant. The ranking  $*_{\text{Ons}}[\eta]$ , ONSET >> DEP-IO(C) >> Linearity-IO accounts for the selection of [uŋ.ŋa] as the output in (23).

(23)/uŋa/ [uŋ.ŋa]

	$*_{\text{Ons}}[\eta]$	ONSET	DEP-IO(C)	Linearity-IO
a. u.ŋa	*!			
b. uŋ.a		*!		
c. uŋ.ŋa			*	*

In order to complete the account of the velarization facts relevant to *unha*, it is necessary to explain the presence of a velar in phrases (postlexically) or when followed by an epenthetic vowel. In other words, why is  $*[alimána]$  ill-formed, given [alimán] and [alimána]? Why are [alimána.to] ‘‘tall German’’ and [alimána] well formed? I claim that an output-to-output constraint requiring identity at the level of the morphological word (stem + terminal elements) is responsible for the selection of velar nasals in vowel-initial phrases, before epenthetic vowels, and in compounds with a vowel-initial second member. IDENT-OOMWd dominates  $*_{\text{Ons}}[\eta]$ , thus producing onset velars postlexically and in compounds, but not in prefixed and suffixed forms. Under IDENT-OOMWd, [alimána] is compared to the output of /alemána/ [alimána] and not to /alemán/ [alimán], since this is a different morphological word. [alimána] is preferred over  $*[alimána]$  because the relevant morphological word to which it is being compared is [alimán]. Since the final vowel is epenthetic, [alimána] must be a prosodic word and thus IDENT-OOMWd is trivially satisfied. In compounds, which are made up of two or more

morphological words, each one of the members is compared to the output of the corresponding word; therefore the ranking IDENT-OOMWd >> \*<sub>Ons</sub>[ŋ] explains forms like [trɛ.ŋa.li.máj]. The constraint IDENT-BA is ranked low (below IDENT-OOMWd) so that identity to the base (for instance, masculine base for the feminine) is violated in [alimá.na].

With regard to the [alimáj]/[alimá.na] alternation, it is useful to consider forms like [irmão] “brother” versus [irmá] “sister” in the central and eastern dialects of Galician. These forms suggest the presence of an empty slot that did not recover the original intervocalic nasal (> Latin GERMANUS > *germano*; Latin GERMANA > *hermana*). One crucial difference, however, is the vowel preceding the nasal. Since this is [a], no feature is available to serve as the point of articulation for a recovered nasal (labial, coronal, dorsal) and thus the intervocalic empty slot is deleted instead. Moaña Galician prefers the final-vowel deletion alternative—[irmáj] “brother” versus [irmá] “sister.” Further evidence in favor of this account can be found in the outcome of nasal deletion (empty slot) after front vowels. As expected, this is a palatal nasal in Galician and Portuguese, for example, *vinho* [bino] “wine.”

Like Galician and Portuguese, Gascon also lost Latin intervocalic nasals. Gascon shares with Galician coda nasal velarization. It differs, however, from both Galician and Portuguese in its tolerance for empty slots. The presence of empty slots in Gascon has been sufficiently demonstrated in Hualde (1992). One piece of evidence presented by Hualde (1992:183) refers to items that historically had initial /h/, but that are now vowel initial, and that trigger the same assimilation process as [h] in other dialects. For instance, the article *eth* /et/ appears as [er] before most vowel-initial words (*eth amic* [eramík] “the friend”); however, with a few words that appear to start with a vowel, there is gemination of the final /t/ (*hum* [ym] “smoke,” *eth hum* [ettym] “the smoke”). These exceptional forms historically had initial aspiration. Hualde convincingly argues that this constitutes solid evidence that in some lexical items a historical /h/ has not been completely lost. Similarly to what I propose here for an earlier stage of Galician, Hualde argues that these forms contain an initial empty consonant slot in their lexical representation, as in (24).

(24)	VC	CVC
	e t	y m

In consonance with the presence of empty slots in Gascon, \*X is low-ranked. The Gascon data also reveal that Linearity-IO dominates IDENT-BA[yŋ], as it is more important to be completely faithful to the UR (including

preservation of precedence relations) than to resemble the masculine form of the indefinite article (25). (For the purpose of the analysis, differences in vowel quality have been ignored. Gascon has a front, round vowel.)

(25) Gascon: ONSET, \*<sub>Ons</sub>[ŋ] >> MAX-IO(Place), IDENT-IO(Nasal) >> Linearity-IO, IDENT-BA[yŋ] >> \*X

(26) Gascon: / $\tilde{y}$ \_a / [ỹ\_a]

	ONSET	* <sub>Ons</sub> [ŋ]	MAX-IO (Place)	IDENT-IO (Nasal)	Linearity-IO	IDENT-BA[yŋ]	*X
a. $\tilde{y}$ ma					**	*!	
b. $\tilde{y}$ _a						*	*
c. $\tilde{y}$ ŋ.a	*!				**		
d. $\tilde{y}$ .ŋa		*!			**		
e. $\tilde{y}$ ŋ.ŋa					**!		

As tableau (26) illustrates, candidates (c) and (d) violate top-ranked constraints; (a) and (e) lose to (b) because they both incur violations of Linearity-IO.

(27) Summary of rankings and effects:

a. *Stage 1*

Galician, Portuguese, Gascon /unam/ [ũ\_a]:

LAZY, MAX-IO >> MAX-IO (Place), IDENT-IO (Nasal) >> Linearity-IO

b. *Stage 2*

Portuguese /ũ\_a/ [ũma], Galician /ũ\_a/ [ũŋ.ŋa], Gascon /ũ\_a/ [ỹ\_a]:

Portuguese: \*X, \*ŋ >> MAX-IO (Place), IDENT-IO (Nasal) >> Linearity-IO

Galician: ONSET, \*X, \*<sub>Ons</sub>[ŋ], IDENT-BA[uŋ] >> MAX-IO (Place),

IDENT-IO (Nasal) >> DEP-IOC, Linearity-IO

Gascon: ONSET, \*<sub>Ons</sub>[ŋ] >> MAX-IO (Place), IDENT-IO (Nasal) >> Linearity

IO, IDENT-BA[yŋ] >> \*X

c. *Stage 3*

Portuguese /ũma/, Galician /uŋa/ [uŋ.ŋa], Gascon /ỹ\_a/

#### 4. Conclusions

This chapter argues that intervocalic velar nasals in *unha* forms in Galician are underlying. In the output, however, underlying velar nasals surface as geminates in order to satisfy syllabic well-formedness constraints. It is argued that an optimality-theoretic framework is necessary to integrate all the interacting factors conspiring to produce the output form. It is also shown that the OT analysis proposed for the Galician data accounts for diachronic and synchronic data in Portuguese and Gascon, shedding light on the grammars of those languages. In addition, the analysis is able to explain the difference

between underlying and derived velar nasals in Galician synchronically and diachronically. Acoustic studies constitute a useful next step in providing evidence for the phonological analysis. The preliminary findings of one study reported by Colina and Díaz-Campos (2003) suggest that underlying and derived velar nasals are acoustically distinct.

#### REFERENCES

- Alvarez, Rosario, X. Luís Regueira, & Henrique Monteagudo. 1986. *Gramática galega*. Galaxia: Vigo.
- Benua, Laura. 1995. "Identity Effects in Morphological Truncation". *Papers in Optimality Theory* ed. by Jill Beckman, Suzanne Urbanczyk, & Laura Walsh, 77-136. Amherst, MA: Graduate Linguistic Student Association.
- Carballo Calero, Ricardo. 1979. *Gramática elemental del gallego común*. Galaxia: Vigo.
- Castro, Obdulia. 1989. *Aproximación a la fonología y morfología gallegas*. Ph.D. dissertation, Georgetown University.
- Clements, George N. 1985. "The Geometry of Phonological Features". *Phonology Yearbook* 2.225-252.
- Colina, Sonia & Manuel Díaz-Campos. 2003. "Syllabic Affiliation of Intervocalic Velar Nasals in Galician". Paper presented at the first Phonetics and Phonology in Iberia Conference, Lisbon, June 2003.
- Dubert, Francisco. 1998. "Reflexións sobre o silabeo de *unha, algunha e ningunha*". *Verba* 25.143-163.
- Hajek, John. 1997. *Universals of Sound Change in Nasalization*. Oxford: Blackwell.
- Hualde, José Ignacio. 1992. "Empty-Consonants and Other CV-Tier Effects in Aranese Gascon". *Lingua* 86:2-3.177-188.
- Ito, Junko & Armin Mester. 1994. "Reflections on CodaCond and Alignment". *Phonology at Santa Cruz* 3.27-46.
- Kaisse, Ellen M. & Patricia Shaw. 1985. "On the Theory of Lexical Phonology". *Phonology Yearbook* 2.1-30.
- Kenstowicz, Michael. 1996. "Base-Identity and Uniform Exponence: Alternatives to cyclicity". *Current Trends in Phonology: Models and methods* ed. by Jacques Durand & Bernard Laks, 363-93. Salford: University of Salford, European Studies Research Institute.
- Kiparsky, Paul. 1985. "Some Consequences of Lexical Phonology". *Phonology Yearbook* 2.85-138.
- Kirchner, Robert. 1998. *An Effort-Based Approach to Consonant Lenition*. Ph.D. dissertation, University of California, Los Angeles.
- Lipski, John M. 1976. "On the Velarization of N in Galician". *Neuphilologische Mitteilungen* 75.182-191.
- , 1994. *Latin American Spanish*. London: Longman.



- Martínez-Gil, Fernando. 1993. "Galician Nasal Velarization as a Case Against Structure-Preservation". *Proceedings of the 19<sup>th</sup> Annual Meeting of the Berkeley Linguistic Society* ed. by Joshua S. Guenter, Barbara A. Kaiser, & Cheryl Zoll, 254-267. Berkeley, CA: Berkeley Linguistic Society.
- , 1997. "Word-Final Epenthesis in Galician". *Issues in the Phonology and Morphology of the Major Iberian Languages* ed. by Fernando Martínez-Gil & Alfonso Morales-Front, 269-342. Washington, DC: Georgetown University Press.
- McCarthy, John. 2002. *A Thematic Guide to Optimality Theory*. New York: Cambridge University Press.
- McCarthy, John & Alan Prince. 1993. "Generalized Alignment". *Yearbook of Morphology* ed. by Geert Booij & Jaap van Marle, 79-153. Dordrecht: Kluwer.
- Meyer-Lübke, Wilhelm. 1921. "The Entwicklung von zwischensilbischem n". *Zeitschrift romanische Philologie* 41.555-65.
- Mohanan, Karuvannur P. 1986. *The Theory of Lexical Phonology*. Dordrecht: Reidel.
- Pater, Joseph. 1999. "Austronesian Vowel Substitution and Other NC Effects". *The Prosody-Morphology Interface* ed. by René Kager, Harry Van der Hulst, & Wim Zonneveld, 310-343. New York: Cambridge University Press.
- Porto Dapena, José Alvaro. 1976. "Fonología de la N velar gallega". *Revista de dialectología y tradiciones populares* 32.467-477.
- Rohlf, Gerhard. 1970. *Le Gascon: Études de philologie pyrénéenne* (Beihefte zur Zeitschrift für Romanische Philologie, 85). Tübingen: Niemeyer.
- Sagey, Elizabeth. 1986. *The Representation of Features and Relations in Nonlinear Phonology*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Veiga, Amable. 1976. *Fonología gallega*. Valencia: Bello.
- Williams, Edwin Bucher. 1962. *From Latin to Portuguese: Historical phonology and morphology of the Portuguese language*. Philadelphia: University of Pennsylvania Press.

# NULL OBJECTS IN FRENCH AND ENGLISH\*

SARAH CUMMINS & YVES ROBERGE  
*Université Laval & University of Toronto*

## 0. *Introduction*

An important subject-object asymmetry in generative grammar has been the obligatory projection of a subject position (by the EPP or a feature of the inflectional layer of the clause) but not of an object position. Projection of an object position was considered to depend on lexical characteristics of the verb. However, languages seem to allow a wide range of possibilities for conventionally intransitive verbs to appear with a direct object (as illustrated for French and English in (1)), and for conventionally transitive verbs to appear without a phonologically realized direct object (2).<sup>1</sup>

- (1) a. *Elle précisa qu'elle le mangerait «tout complètement», feula des baisers à blanc et raccrocha.* (L:110)  
“She added that she would eat him ‘all up,’ growled air kisses, and hung up.”
- b. *Si Mike commence à bafouiller ses tirs, la sauce commence à prendre avec ses partenaires.* (L:113)  
“While Mike is beginning to splutter his shots, things are coming together for his teammates.”
- c. Just how far the argument has come since Archie bellowed his brand of bigotry is evident in the first episode of 704 Hauser Street.  
(*Montreal Gazette* 19 December 1993)
- d. Two young German women wept tears of shame for their country as the car left.  
(*Ottawa Citizen* 4 June 1993)
- (2) a. *La lune, si t’y mets une porte et tu regardes \_\_ la nuit, tu peux être fier de ton boulot.* (Gourio 2002:153)  
“If you put a door on the moon and you watch \_\_ at night, you can be proud of your work.”

---

\* We would like to thank Denis Bouchard, Diane Massam, Philippe Prévost, Michelle Troberg, Barbara Vance, three anonymous reviewers, and the members of the Asymmetry Project. This work is supported by a grant from the Social Sciences and Humanities Research Council of Canada (Di Sciullo 412-97-0016).

<sup>1</sup> Abbreviations in examples: L: Larjavaara (2000); GV & PM: García Velasco and Portero Muñoz (2002); BNC: British National Corpus (general corpus).

- b. *C'est pas lui qui l'a écrit, son livre, le pape, c'est quelqu'un qui lui écrit* \_\_.  
 (Gourio 2002:153)  
 "The Pope didn't write his book himself, someone writes \_\_ for him."
- c. Why then do the psychic gifts often seem to tease \_\_, confuse \_\_ and obstruct \_\_?  
 (BNC:B2G 1620)

These possibilities cannot be attributed solely to lexical properties of the verb; if this were the case, certain verbs would always be able to appear without their objects regardless of the construction or discourse context, and others would never be able to appear without an object. As we will show, this is not the case. Rather, following Roberge (2003), we propose that null or implicit objects can be attributed to a Transitivity Requirement (TR) just as null subjects are ultimately due to the EPP. Recoverability for the EPP is morphologically based, as is evident in null-subject languages, while recoverability involving the TR may also be semantically and pragmatically based; as we will show later, such recovery may be based on information derived from the verb's lexical semantics and generalized conversational implicatures (formalized as in Levinson 2000) involved in the interpretation of reduced nominal forms. The factors that contribute to licensing superficial intransitivity—the absence of an overt object—may include lexical semantics, functional elements, discourse factors, and transclausal structural elements. This view is supported by a comparative study of null-object possibilities in French and English.

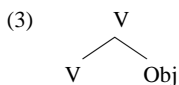
### 1. *On transitivity*

The concept of transitivity has been interpreted as a continuum in certain works, and a distinction has been proposed between syntactic transitivity and semantic transitivity; see, among many others, Blinkenberg (1960), Desclés (1998), Hopper and Thompson (1980), and Lazard (1994). Surprisingly little is ever said about the object position itself. The hypothesis in Roberge (2003) is that there exists a Transitivity Requirement (TR), whereby an object position is always included in VP, independently of the lexical choice of V. The empirical motivation of this hypothesis is the well-documented evidence (see in particular Blinkenberg; Larjavaara 2000) that virtually any transitive verb has the potential to appear without a direct object and virtually any unergative verb has the potential to appear with a direct object, given an appropriate context.<sup>2</sup>

---

<sup>2</sup> One context that enhances the acceptability of null objects is a habitual present; see example (10a), with *devour* (often cited as a verb unable to appear without an overt object). In our judgment, virtually any English verb could appear felicitously in a sentence analogous to this one.

To account for these facts, there must be a mechanism to generate the direct-object position, either optionally or obligatorily. The TR represents the second, more restrictive, possibility and conveys the concept of transitivity as a property of the predicate (the VP), rather than as a property of the lexical content of V. The TR is the internal-argument counterpart to the EPP. In other words, the configuration in (3)—order irrelevant—is given by UG:



The TR is similar to the EPP in that, first, it targets a position and not necessarily the nature of the element occupying this position and, second, the end result varies depending on lexical choice and the merger and movement operations involved in the derivation. It differs from the EPP in that, first, it does not target a Spec position and, second, it is active in the thematic layer of the clause rather than the inflectional layer. For the purpose of our discussion, we define the null object interpretatively as an  $x$  such that  $x$  is: phonologically null, involved in the event denoted by the VP, and not an external argument.<sup>3</sup>

## 2. *Toward a typology of null objects*

Two recent studies—Larjavaara (2000) on French and García Velasco and Portero Muñoz (2002) on English—address the issue of null objects comprehensively, while taking account of previous work on this topic. The findings of these two studies show clear similarities between the two languages. Both studies distinguish two types of objects: García Velasco and Portero Muñoz call the two types indefinite and definite null objects, while Larjavaara refers to generic and latent null objects. Examples of the two types are illustrated in (4) and (5):

- (4) Indefinite/generic:
- a. Do you write \_\_? (GV & PM:4)
  - b. *Wild Guns est un jeu qui défoule* \_\_ . (L:88)  
“Wild Guns is a game that destresses \_\_.”

---

<sup>3</sup> Note that this definition correctly excludes empty object positions that are directly linked to an element in external argument position such as in passives, unaccusatives, and perhaps middles. However, it leaves open to a null object interpretation an eventual unexpressed object position in unergative VPs. The definition also includes null oblique objects, although we will not discuss them here.

- (5) Definite/latent:
- a. “Do you like \_\_\_?” “I love \_\_\_!” (GV & PM:12)
  - b. «*Tu as lu les pages?*» *Il avait lu* \_\_\_. (L:43)  
“‘Did you read the pages?’ He had read \_\_\_.”

Both studies note characteristics of one or the other type. García Velasco and Portero Muñoz (2002) point out that definite objects are typically a non-first-order entity; Larjavaara (2000) notes that the latent object often has propositional content. The two agree that indefinite or generic null objects do not have a contextually available referent. García Velasco and Portero Muñoz point out that generic null objects can give rise to an activity rather than an accomplishment reading of the verb; Larjavaara notes that null objects can focus attention on the activity. Both point out that the lexical characteristics of the verb can help to identify the referent of the null object. García Velasco and Portero Muñoz note that null objects are often found in fixed phrases, while Larjavaara describes a wider context of de-actualization as being favorable to null objects. And both note several structural contexts that favor a nonovert object. These contexts are summarized and illustrated in (6) through (12).

- (6) Sequences of verbs:
- a. He will steal \_\_, rob \_\_, and murder \_\_. (GV & PM:2)
  - b. *Elles ont caressé* \_\_, *pétri* \_\_, *étreint* \_\_, *pénétré* \_\_. (L:97)  
“They have caressed \_\_, kneaded \_\_, clasped \_\_, penetrated \_\_.”
- (7) Imperatives:
- a. Push \_\_ hard. (GV & PM:2)
  - b. *Fais voir* \_\_. (L:50)  
“Show \_\_.”
- (8) Contrastive uses:
- a. He theorizes about language, but I just describe \_\_. (GV & PM:2)
  - b. *Seulement moi, je n’assassine pas* \_\_, *je ressuscite* \_\_. (L:91)  
“Only I don’t murder \_\_, I resuscitate \_\_.”
- (9) Infinitive:
- a. This is a lovely guitar, with an uncanny ability to impress \_\_ and delight \_\_. (BNC:C9K 1330)
  - b. *Pour compenser* \_\_, *j’ai décidé d’adopter dorénavant cette graphie*. (L:85)  
“To compensate \_\_, I have decided to use that spelling from now on.”
- (10) Generic present tense:
- a. There are those who annihilate \_\_ with violence—who devour \_\_. (BNC:FAT 2709)
  - b. *Un peintre dérange* \_\_ *bien moins qu’un écrivain*. (L:83)  
“A painter disturbs \_\_ much less than a writer.”
- (11) Dative pronoun (French):
- J’étais où quand tu lui avais donné* \_\_? (L:39)  
“Where was I when you gave \_\_ to him?”

(12) *Ça* as subject (French):

*Ça flingue \_\_ à tout va là-dedans.* (L:91)  
 “They’re shooting \_\_ like crazy in there.”

In a third study, Goldberg (2001) investigates unexpressed objects of causative verbs (those that entail a change of state in the patient argument) in English. She concludes that the option of leaving these arguments unexpressed depends largely on factors relating to information structure: The unexpressed object is typically neither topical nor focal, and the verb is emphasized somehow, by being iterative or generic, by being contrasted with another verb, or by having a narrow focus.

All of these authors implicitly or explicitly adopt the position that the missing argument is not syntactically represented: Syntactically the verb is intransitive. In a generative framework, this position finds a counterpart in Rizzi (1986:509-510), who proposes that both the arbitrary third-person human interpretation, meaning “people in general” or “some people,” and the prototypical-object interpretation, where the verb’s lexical semantics identify the object, are available lexically to saturate the argument’s theta role and block projection. Thus, the verbs are intransitive in syntax. The absence of a syntactic object explains why, in Rizzi’s account, the type of sentence exemplified in (13) is impossible in English: There is no object that can bind the anaphor or be modified by the adjective. However, such sentences are grammatical in Romance; hence several accounts (Rizzi 1986; Authier 1989; Roberge 1991) posit a syntactically present null object.

(13)a. *Ce gouvernement rend \_\_ malheureux.*

\*“This government makes \_\_ unhappy.”

b. *Une bonne bière reconcilie \_\_ avec soi-même.*

\*“A good beer reconciles \_\_ with oneself.”

Under the TR, the object position is projected and the verb remains transitive in syntax in both English and French. Although we do not find sentences like those in (13) in English (as shown by the ungrammaticality of the glosses), there is nonetheless evidence that a null object has an effect on syntax in both English and French.<sup>4</sup> For example, null objects can enter into a

---

<sup>4</sup> The TR redefines the notion of null object and broadens the range of phenomenon it subsumes. It is beyond the scope of this chapter to deal with all of them; we will not explore here the structure exemplified in (13) or attempt to explain the differences it highlights between Romance languages and English. Moreover, we do not investigate null objects with clausal or propositional characteristics.

network of relationships with compatible pronouns, and sometimes require coreference, either with pronouns or with another null object, as in (14):

- (14) a. *Ce roman amuse \_\_\_ quand on le prend avec humour.*  
 “This novel amuses \_\_\_ if one takes it with a sense of humor.”  
 b. *Qui aime\_\_ bien châtie \_\_ bien.*  
 “Who loves \_\_\_ well punishes \_\_\_ well.”  
 c. His attitude intimidates \_\_\_, until you figure out he’s a phony.  
 d. It’s better to reuse \_\_\_ than to recycle \_\_\_.

Null objects can serve as the argument of a secondary predicate, as in (15):

- (15) a. *Les steaks, moi, je préfère manger \_\_\_ saignant.*  
 “Steaks, I like to eat \_\_\_ rare.”  
 b. *Vous avez acheté \_\_\_ en solde?*  
 “Did you buy \_\_\_ on sale?”  
 c. Beat \_\_\_ until thick and lemon-colored.

A syntactically represented null object is required to account for the availability of a parasitic-gap interpretation for sentences such as (16).

- (16) Which document did the spy memorize\_\_ before eating \_\_\_?

This shows the necessity, even under a lexical account, of projecting an empty argument position. A lexical account, moreover, would require three mapping patterns for verbs such as *eat*: transitive with overt object, transitive with null object, and intransitive, for the prototypical-object or activity reading.

Finally, a null object can receive further specification. Example (17a) shows further specification of the null object of a transitive verb, while (17b) shows an attempt to further specify the argument of an unaccusative. The result is uninterpretable, presumably because the argument has moved. The impulse is to try to interpret (17b) as a transitive—to supply a null object.

- (17) a. *C’est une chose si douce que de louer \_\_\_, et surtout ses amis.* (L:82)  
 “To praise \_\_\_ is such a sweet thing, and especially one’s friends.”  
 b. *\*C’est une chose si difficile que de partir, et surtout ses amis.*  
 “To depart is such a difficult thing, and especially one’s friends.”

These facts argue against both the lexical and the constructional accounts, which treat such sentences as objectless. Rizzi’s (1986) general discussion also leaves unexplained instances of null objects that receive neither the arbitrary-human nor the prototypical-object interpretation, such as those in (18). Instead,

elements of the linguistic and extralinguistic contexts come into play. It seems obvious that such information is not part of the lexical entry of the verb.

- (18) a. Lifting his arm to strike \_\_\_\_, he felt a grip of iron around his wrist, restraining him. (BNC:ACV 1183)  
 b. *On voit que ce n'est pas lui qui lave \_\_\_\_.* (L:86)  
 "You can tell he's not the one who washes \_\_\_\_."  
 c. *M. Jospin, maintenant, régularisez \_\_.* (L:55)  
 "Mr. Jospin, now, regularize." (protesters' banner referring to the situation of immigrants without papers)  
 d. When you don't have money and you have to work hard to accomplish \_\_\_\_ in life, it's not that easy to just throw it down... (Houpt 2003)

Moreover, if the absence of an overt object could be explained entirely in semantic and pragmatic terms, we would expect English and French null objects to be substantially the same. But in fact, there is a subset of Larjavaara's (2000) latent objects in French that have no counterpart in English. Examples are shown in (19). In these cases, there is a specific linguistic referent in the context, and the only interpretation is that this antecedent is the referent of the null object.

- (19) a. *On lui tendit une main... Vexé, il négligea \_\_.* (L:48)  
 "A hand was extended to him. Annoyed, \*he ignored \_\_\_\_."  
 b. *Si un mec t'offre un café balance \_\_ lui à travers la gueule.* (L:50)  
 "If a guy offers you a coffee, \*throw \_\_\_\_ in his face."  
 c. *Nikel m'a dit de prendre une boîte bleue dans le vestiaire. J\_' ai prise \_\_.* (L:59)  
 "Nikel told me to take a blue box from the locker. \*I took \_\_\_\_."

These absent objects, which are taken as definite and referential, resemble null arguments discussed by Huang (1984), Farrell (1990), and Cardinaletti (1990), among others, and analyzed as variables bound by a null topic or as null pronouns. In either case, the object is taken to be syntactically present. This is the position we adopt for the full range of null objects in French and English, by virtue of the TR, and we turn now to the issue of how these null objects are licensed and recovered.

### 3. *Recoverability of null objects*

Null objects are diverse, and so are the means of their recovery. We propose that there are three means of recovering the identity or reference of null objects: internally, through material in IP; through discourse, involving referential null objects; and by binding from the left periphery, that is, by a topic. We take up each of these in turn.



### 3.1 Internally licensed null objects

All of García Velasco and Portero Muñoz' (2002) "understood" objects, all of Goldberg's (2001) "omitted" arguments, all of Larjavaara's (2000) generic "absent" objects and many of her latent "absent" objects can be considered to be internally licensed, recovered through material in the IP. A primary means of recovery comes from lexical characteristics of the verb, as with the true prototypical-object interpretation. Note that the prototypical object of psychological verbs, which are commonly found with null objects in both English and French, is in fact the arbitrary third-person affected-human interpretation (see (13)).

- (20) a. *La magie des séries, c'est de surprendre \_\_, de dépayser \_\_.* (L:98)  
 "The magic of the playoffs is in surprising \_\_, disorienting \_\_."  
 b. Where Boulestin never falters or misleads \_\_ is in the sureness of his taste and the sobriety of his ingredients. (BNC:EFU 223)  
 c. ...the patter of the camp, grey-haired one between songs can irritate \_\_.  
 (BNC:A4A 234)

The identity of lexically determined null objects can range from the vaguely predictable, as in (21a) (the area around me); to the narrowly determined, as in (21b) (a paper or envelope); to the entirely predictable, as in (21c)—semantically, the only possible object of *déciller* is eyes. Examples (21a) and (21b) thus illustrate how the lexical-semantic contribution from the verb may be augmented by information from the linguistic and extralinguistic contexts, while (21c) shows an entirely lexical contribution.

- (21) a. «*Ben, qu'est-ce que tu fais?*» *J'explore \_\_.* (L:83)  
 "“Hey, what are you doing?” I'm exploring \_\_.”  
 b. *Dans ma hâte à décacheter \_\_, j'ai déchiré la feuille.* (L:76)  
 "In my haste to unseal \_\_, I tore the page.”  
 c. *Crystal claquait dans ses mains. On décilla \_\_.* (L:54)  
 "Crystal clapped his hands. We opened \_\_.”

The internally licensed null object is not formally linked to another linguistic element. It does not refer; it is not an anaphor and it is not in a relationship with a [+specific] nominal; in Larjavaara's (2000) and García Velasco and Portero Muñoz' (2002) canonical cases, moreover, there is no contextually available referent. In English, when a referential interpretation is forced, a null object is impossible, as in (22), while in a similar context but without forcing reference, the null object is fine, as in (23).

- (22) a. What happened to that carrot?  
 \*I chopped \_\_\_\_ . (Goldberg 2001:512)
- b. The door is open. \*Didn't you lock \_\_\_\_?
- (23) a. What happened to all the vegetables?  
 Well, Jacques has been chopping \_\_\_\_ and dicing \_\_ all afternoon.
- b. Did you lock \_\_? (pulling out of the driveway)

Because the internally licensed null object does not refer and is not anaphoric, pragmatics has a free hand in interpretation, and contextual factors can contribute to the inference of a specific reference. In fact, according to Levinson's I-principle (2000:114), based on Grice's (1975) maxim of informativeness, hearers will seek out a maximally pertinent interpretation of such null objects, assuming rich connections with contextual information. This is illustrated by the sentences in (24).

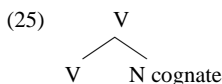
- (24) a. We have to get rid of all the ugly dishes before your date arrives.  
 Okay, you wash \_\_\_\_ and I'll dry \_\_\_\_ . (Goldberg 2001:515)
- b. *Allez, envoie \_\_* . (L:50)  
 "Come on, hand \_\_ over."
- c. I'll introduce \_\_ . (one host to another before a talk)
- d. *Même avec trois cuillerées de sucre en poudre, le breuvage reste amer. Leroy touille \_\_ en comptant les miettes sur la toile cirée* . (L:49)  
 "Even with three spoonfuls of sugar, the drink still tastes bitter. Leroy stirs \_\_, counting the crumbs on the oilcloth."

Other factors that enhance recoverability are found within IP. These include the factors that contribute to de-actualization, such as the generic present tense, the infinitive, and *ça* as subject (see (6)-(12)). Tenseless verb forms and nonreferential tenses favor a nonreferential reading, while referential tenses, such as perfectives, favor a specific, referential reading. Although the correspondence is not perfect in either English or French (nonreferential null objects are attested in sentences with, e.g., perfective tenses), there is a clear tendency to associate specific, referential entities with specific events set at a specific time; for this reason, nonreferential null objects can be less felicitous with referential tenses.

The internally licensed null object can be described as a 'null cognate object.' Overt cognate objects, if unmodified, add no semantic information beyond that contained in the verb itself. The null cognate object is similar, and that is why constructions with null cognate objects are described as focusing on the action or on the verb. We liken the interpretation of a predicate containing a null cognate object to thethetic interpretation: "An assertion is being made as to the existence of an object or of an event involving the object"

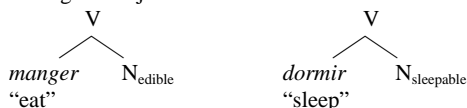
(Basilico 1998:542). But the object is not singled out from the event for a second judgment, such as assignment of a property, as is the case in the categorical judgment.

We propose that the null cognate object is structurally a bare empty noun, similar to the empty NUMBER noun proposed by Kayne (2002). Kayne argues, on the basis of the adjective-like properties of *few* (comparative and superlative forms, distributional facts), that *few* is in fact an adjective modifying a phonologically empty N with the semantic content NUMBER. Null cognate objects can be conceived of as a similar N whose semantic content is derived from the verb, thus one that is semantically cognate to the verb as in (25):

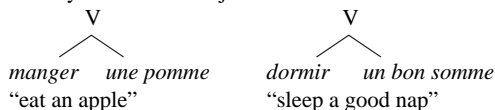


The null cognate object is available for all verbs. The difference between conventionally transitive verbs (such as *manger* “eat”) and ‘unergatives’ (such as *dormir* “sleep”) is that the null cognate object is the more marked object for the former class and the less marked object for the latter class (26a-b). Moreover, both classes can have objects that are semantically independent of the verb, as in (26c).

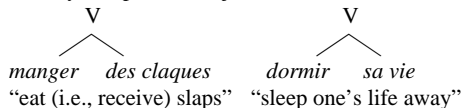
(26) a. Null cognate object:



b. Lexically conditioned object:



c. Lexically independent object :



### 3.2 Referential null objects

Certain null objects, like those in (19) and (27), have a referent that is identifiable from the linguistic or extralinguistic context. In this, they differ

sharply from the internally licensed null objects discussed in the previous section.

- (27) a. «*Maîtrisez-vous vos interviews? C'est capital, les interviews.*» *Je maîtrise* \_\_\_\_ .  
(L:50)  
“‘Do you control your interviews? Interviews are very important.’ \*I control \_\_\_\_.”
- b. *Ça provenait de deux planches à dessin dressées en guitoune. Crystal contourna* \_\_\_\_ . (L:88)  
“It came from two drawing tables used as a tent. \*Crystal bypassed \_\_\_\_.”

These null objects appear similar to null objects of colloquial German, discussed by Cardinaletti (1990) and exemplified in (28):

- (28) *Habe ich* \_\_\_\_ *gestern gekauft.*  
have I yesterday bought  
“I bought \_\_\_\_ yesterday.”

Under Cardinaletti's analysis these null objects involve a base-generated empty operator locally binding a null pronominal variable; the content of pro is determined by the operator, which is compatible with third-person pronouns only. Kampen (1997) discusses similar examples in informal Dutch and presents evidence that the null topic is a null pronoun that is underspecified with respect to phi-features. Crucially, null pronominal variables may not appear if SpecCP is filled by lexical material as in (29):

- (29) a. \**Gestern habe ich* \_\_\_\_ *gesehen.*  
yesterday have I seen
- b. \**Wann hast du* \_\_\_\_ *gesehen?*  
when have you seen

The null objects of French in (19) and (27) do not fall under the same constraints. Although third-person reference is most common, it is possible to construct acceptable examples with second- or first-person reference as in (30):

- (30) a. *Crystal tient à toi. Mais les méchants veulent* \_\_\_\_ *lui prendre.*  
“Crystal wants to keep you. But the bad guys want to take \_\_\_\_ from him.”
- b. *Yan m'a vue pour la première fois à la bibliothèque, et tout de suite il a adoré* \_\_\_\_!  
“Yan saw me for the first time in the library, and right away he adored \_\_\_\_!”

Moreover, the presence of lexical material in SpecCP does not prevent the appearance of a null object.

- (31) a. *Si on prenait Tigre et Dragon? Qui a vu \_\_\_?* (in video store)  
 “How about *Crouching Tiger Hidden Dragon*? Who has seen \_\_\_?”  
 b. *Tu as lu les pages? Tu m’as dit que tu avais lu \_\_\_.*  
 “Did you read the pages? You told me you had read \_\_\_.”

Therefore, there is no evidence for an operator in CP influencing the empty object in French.<sup>5</sup> Rather, this type of null object seems to correspond to a clitic, and its appearance can be seen as an instance of clitic drop. The simplest approach to this construction would be to assume that it corresponds exactly to its counterpart with an accusative clitic; see Tuller (2000), Guasti and Cardinaletti (2003:fn. 17) and references cited therein. Semantically, the sentences are equivalent to corresponding sentences with an object clitic, and the structural contexts for this type of null object are identical to those of clitics; see (32).

- (32) a. A: *Pourquoi avoir choisi cette époque?*  
 “Why did you choose that period?”  
 B: *Parce que j’adore \_\_\_.* (L:64) = *Parce que je l’adore.*  
 “Because \*I love \_\_\_/I love it.”  
 b. *Nikel m’a dit de prendre une boîte bleue dans le vestiaire. J’ai prise \_\_\_.* = *Je l’ai prise.* (L:77)  
 “Nikel told me to take a blue box from the locker. \*I took \_\_\_/I took it.”

Moreover, the attested past-participle agreement in (32b), while certainly unusual, is identical to the agreement that would be found if an accusative clitic were present. Thus, *pro* can appear without the clitic that is normally used to recover its feature contents. We do not address the issue of determining whether an empty or silent clitic linked to *pro* must be postulated or whether *pro* can appear on its own with default features. The function of object clitics is to morphologically recover definite null objects; it seems that in these sentences, this recovery mechanism is manipulated for stylistic effect. English has no similar element able to recover null objects; therefore this option is not available and all counterparts to the clitic-drop examples are ungrammatical in English.

<sup>5</sup> A similar conclusion (i.e., that the null object in (27) is not a null variable) would be reached through an application of the tests proposed in Raposo (1986). For example, when talking about a safe, it would be acceptable in French to say:

(i) *J’ai informé la police de la possibilité que la secrétaire ait ouvert \_ à l’insu de son patron.*  
 “I informed the police of the possibility that the secretary might have opened without her boss’s knowledge.”

Raposo uses the ungrammaticality in European Portuguese of a similar construction to argue that this type of null object is a variable in this language.

3.3 Topics in English and French

French and English, as is well known, display contrasting behavior of topics and the linked element in the matrix. Topics of the type in (33) are linked directly to an empty object in the matrix in English. In French, the link is mediated by a clitic as in (34), in the structure dubbed Clitic Left Dislocation by Cinque (1990).

- (33) a. Your book, I bought \_\_. / \**Ton livre, j'ai acheté* \_\_.
- b. John, I can't stand \_\_. / \**Jean, je ne supporte pas* \_\_.
- (34) a. *Ton livre, je l'ai acheté* \_\_.
- b. *Jean, je ne le supporte pas* \_\_.

Rizzi (1997) posits [Your book OP I bought *ec*] for the English case in (33a), where the *ec* is a null constant licensed by the anaphoric operator. In Romance, according to Rizzi, clitics fill the same function of establishing the connection between the topic and the open position in the comment. We adopt this analysis, but note that the facts regarding topics, clitics, and *ecs* are somewhat more complex. Taking indefinite topics into account and incorporating the discourse-linked *ecs* discussed in section 3.2 into the mix, the similarities and contrasts between English and French emerge as summarized by the data in (35) and (36).

(35) Overt topic

	a. Token	b. Type
English	That book, I hated __. John, I can't stand __.	Wine, I bought __. Bananas, I'll eat __.
French	* <i>Ce livre, j'ai détesté</i> __. * <i>Jean, je ne supporte pas</i> .	<i>Vin, j'ai acheté</i> __. (cf. also <i>Vin, j'en ai acheté</i> .) <i>Les bananes, je mange</i> __. (cf. also <i>Les bananes, j'en mange</i> .)

(36) No topic or null topic

	a. Definite (linguistic antecedent)	b. Indefinite (no linguistic antecedent)
English	Did you read the pages? *He had read __. A: What do you think of my cake? B: *I like __.	A: You like __ ? B: I love __ ! ? So, how would you rate __ ?
French	<i>Tu as lu les pages?</i> <i>Il avait lu</i> __. A: <i>Que penses-tu de mon gâteau?</i> B: <i>J'aime</i> __ !	A: <i>Tu aimes</i> __ ? B: <i>J'adore</i> __ ! <i>Alors, comment as-tu trouvé</i> __ ?

In summary, nondefinite null objects are acceptable in both languages, as in (35) and (36b); they are linked to an overt topic or an element in the

nonlinguistic context. In English, a definite null object must be linked to an overt topic; compare (35a) and (36a).

In French, a definite token null object is typically linked to an overt topic only via a clitic (35a). However, there are innovative instances of a definite token topic and a null object linked without an overt clitic (see Fonágy 1985). In Fonágy's corpus, these cases most commonly involve the verbs *connaître* or *aimer* but are not restricted to these.

- (37) a. *Jacques F., vous connaissez \_\_?* (Fonágy 1985:5)  
 "Jacques F., do you know \_\_?"  
 b. *Le yogourt X, il aime \_\_, il adore \_\_.* (Fonágy 1985:8)  
 "X yogurt, he loves \_\_, he adores \_\_."  
 c. *La bleue, je prends \_\_.* (Fonágy 1985:9)  
 "The blue one, I'll take \_\_."

These cases show clear parallels with the clitic-drop cases discussed in section 3.2 and represented in (36a). They are semantically equivalent to corresponding sentences with a clitic, and the sole structural difference is the absence of a clitic-linking topic and null object. We hypothesize that clitic drop is the strategy at work with the overt topics, as well as in the cases involving a null topic or no topic. Clitic drop can be seen as an extension of the general pragmatic strategy involved in the interpretation of nominal elements. We assume that a standard pattern in discourse is the sequence [lexical noun...pronoun...null object] in which all nominals are interpreted as coreferential (assuming no contradictory information). This is exemplified in (38):

- (38) *J'ai vu ton chien<sub>1</sub> dans le parc. Je l<sub>1</sub>'ai caressé Ø<sub>1</sub>.*  
 "I saw your dog in the park. I petted it."

The coreference between the pronoun and the null object is established by purely grammatical means, while that between the lexical noun and the pronoun is pragmatic and defeasible, involving further application of Levinson's (2000) I-principle, whereby a hearer infers from a lack of specification that there is no need for specification. The default here is for the hearer to assume coreference between the lexical noun and the pronoun. With clitic drop, the same implicature comes into play, this time between the lexical noun and the null object.

Unlike null cognate objects, the instances of clitic drop are considered innovative and stylistically marked. Fonágy (1985) drew most of his examples from younger speakers and from advertising and considers that, in the latter

case at least, the strategy is a deliberate attempt to appear hip and fresh. Larjavaara (2000) limited her corpus to recent works and chose literary texts that aim for a style that could be described in the same terms. It goes without saying that many such examples will be considered ungrammatical by speakers with other demographic profiles and different stylistic aims. However, as Lambrecht and Lemoine (1996:280) remark, data from spoken varieties must be included in a complete account of grammatical phenomena, especially in a language, like French, whose written variety is greatly influenced by prescriptive influences and differs significantly from spoken varieties.

In addition, the label ‘grammatical’ is of limited usefulness in assessing the stylistic effect of the strategy of clitic drop. We hypothesize, as a matter for further research, that the stylistic effects noted by Fonágy (1985) and Larjavaara (2000) are due in part to the fact that the result of clitic drop is identical to a null object whose referent is physically salient in the discourse. Noailly (1997) characterizes this use as ‘deictic’ and describes its function as lending cohesion to the discourse. If null objects resulting from clitic drop are similar, it is easier to begin to understand the immediacy they bring to the discourse and their effect of engaging the reader or hearer.<sup>6</sup>

Fonágy (1985) and others view the null objects resulting from clitic drop as a fairly recent phenomenon in French. (We note also that most native speakers we have consulted agree that they are more typical of European French than of Canadian French.) However, Arteaga (1998) argues for a syntactically present null object in Old French, which she analyzes as a null pronominal. She identifies three contexts for null objects: left-dislocation structures, like the topic structures in (37); ‘*écrasement*’ structures, comparable to (11); and coordinated structures, as in (39). We note also example (40), a topic construction, from a 14<sup>th</sup>-century text of Middle French (Troberg 2004:8):

- (39) *Il retrait s'espee et met ou fuerre.* (Arteaga 1998's example (3))  
 = *Il retire son épée et \_\_ met \_\_ au feu.* (Modern French)  
 “He pulls back his sword and puts \_\_ in the fire.”
- (40) *Car les lettres que li messages apportoit, c'estoit mes usages de \_\_ regarder \_\_ avant toute oeuvre.* = *Les lettres que le messenger apportait, c'était mon habitude de \_\_ regarder \_\_ avant tout autre travail.* (Modern French)  
 “The letters the messenger brought, it was my habit to \_\_ look at \_\_ before any other work.”

---

<sup>6</sup> Since English does not have the strategy of clitic drop, these most stylistically marked examples of French (e.g., (27)) are completely unacceptable. The stylistic effects of ‘deictic’ uses in English appear similar to their effects in French; the cohesion created between discourse and context marks them as conversational and informal.



Further research is thus needed to establish clearly whether clitic drop is an innovation in contemporary French.

#### 4. *Conclusion*

We have argued that the existence of null objects is largely determined by the TR and that cross-linguistic variation is therefore predicted to occur mostly in the recoverability mechanisms particular grammars use. A comparison of English and French null objects, drawing on corpus data, was used to support this claim. The data from these two languages lead to the conclusion that there are three types of null objects:

1. Bound: a bound variable or a null constant
2. Discourse-linked: a null pronominal
3. Internally licensed: null cognate objects (predicted by the TR), a bare N

It was shown that French and English differ only in the availability of discourse-linked null objects and bound definite null objects (see (35) and (36) for a summary). We have characterized the French discourse-linked null object as a clitic-drop construction.

This preliminary account is intended to sketch out the semantic and syntactic characteristics of the three types of null objects, and the role pragmatic principles play in their recovery.<sup>7</sup> Under the TR, all null objects are syntactically represented; ensuring syntactic representation allows for an account of differences in referentiality and syntactic activity.

#### REFERENCES

- Arteaga, Deborah. 1998. "On Null Objects in Old French". *Romance Linguistics: Theoretical perspectives* ed. by Armin Schwegler, Bernard Tranel, & Myriam Uribe-Etxebarria, 1-12. Amsterdam & Philadelphia: John Benjamins.
- Authier, J.-Marc. 1989. "Arbitrary Null Objects and Unselective Binding". *The Null Subject Parameter* ed. by Osvaldo Jaeggli & Ken Safir, 69-109. Dordrecht: Reidel.

---

<sup>7</sup> Further work is needed on: first, the distinction between nonreferential yet not generic null cognate objects (often with inferred reference to a salient entity in the non-linguistic context) and referential null objects resulting from clitic drop; and second, the issue of null objects in the process of language change and grammaticalization. Referential null objects seem to be gaining ground in spoken French; is this a sign of a shift from the status of a 'syntactic' language to that of a 'pragmatic' language (Huang 2000:261)? Use of null anaphora, like reliance on a topic-comment structure, also increasingly prominent in spoken French, is associated with the latter type.

- Basilico, David. 1998. "Object Position and Predication Forms". *Natural Language and Linguistic Theory* 16.541-595.
- Blinkenberg, Andreas. 1960. *Le problème de la transitivité en français moderne: essai syntacto-sémantique*. Copenhagen: Historisk-filosofiske Meddelelser.
- British National Corpus. <http://www.natcorp.ox.ac.uk> (July-August 2002).
- Cardinaletti, Anna. 1990. "Subject/Object Asymmetries in German Null-Topic Constructions and the Status of SpecCP". *Grammar in Progress* ed. by Joan Mascaró & Marina Nespor, 75-84. Dordrecht: Foris.
- Cinque, Guglielmo. 1990. *Types of A' Dependencies*. Cambridge, MA: MIT Press.
- Desclés, Jean-Pierre. 1998. "Transitivité sémantique, transitivité syntaxique". *La transitivité* ed. by André Rousseau, 161-180. Villeneuve-d'Ascq: Presses Universitaires du Septentrion.
- Farrell, Patrick. 1990. "Null Objects in Brazilian Portuguese". *Natural Language and Linguistic Theory* 8.325-346.
- Fonagy, Ivan. 1985 "J'aime □. Je connais □. Verbes transitifs à objet latent". *Revue Romane* 20.3-35
- García Velasco, Daniel & Carmen Portero Muñoz. 2002. *Understood Objects in Functional Grammar* (Working Papers in Functional Grammar 76). Amsterdam: Functional Grammar Foundation (University of Amsterdam).
- Goldberg, Adele. 2001. "Patient Arguments of Causative Verbs Can Be Omitted: The role of information structure in argument distribution". *Language Science* 23.503-524.
- Gourio, Jean-Marie. 2002. *Brèves de comptoir*. Paris: Laffont.
- Grice, H. Paul. 1975. "Logic and Conversation". *Syntax and Semantics, Vol. 3, Speech Acts* ed. by Peter Cole & Jerry L. Morgan, 41-58. New York: Academic Press.
- Guasti, Maria Teresa & Anna Cardinaletti. 2003. "Relative Clause Formation in Romance Child's Production". *Probus* 15.47-89.
- Hopper, Paul J. & Sandra A. Thompson. 1980. "Transitivity in Grammar and Discourse". *Language* 56.253-299.
- Houpt, Simon. 2003. "Bigotry for Bucks". *Toronto Globe and Mail*, 8 March. (Available online at: <http://www.globeandmail.com/servlet/story/RTGAM.20030307/whillbilly0702/BNStory/Front/>)
- Huang, James. 1984. "On the Distribution and Reference of Empty Categories". *Linguistic Inquiry* 15.531-575.
- Huang, Yan. 2000. *Anaphora. A Cross-linguistic Approach*. Oxford: Oxford University Press.
- Kampen, Jacqueline van. 1997. *First Steps in Wh-Movement*. Delft: Eburon.
- Kayne, Richard. 2002. "On the Syntax of Quantity in English". Ms., New York University.
- Lambrech, Knud & Kevin Lemoine. 1996. "Vers une grammaire des compléments zéro en français parlé". *Absence de marques et représentation de l'absence I* ed. by Jean Chuquet & Marc Frid, 279-309. Rennes: Presses Universitaires de Rennes.
- Larjajaara, Meri. 2000. *Présence ou absence de l'objet. Limites du possible en français contemporain*. Helsinki: Academia Scientiarum Fennica.
- Lazard, Gilbert. 1994. *L'actance*. Paris: Presses Universitaires de France.
- Levinson, Stephen C. 2000. *Presumptive Meaning: The theory of generalized conversational implicature*. Cambridge, MA: MIT Press.
- Noailly, Michèle. 1997. "Les mystères de la transitivité invisible". *Langages* 127.96-109.
- Raposo, Eduardo. 1986. "On the Null Object in European Portuguese". *Studies in Romance Linguistics* ed. by Osvaldo Jaeggli & Carmen Silva-Corvalan, 373-390. Dordrecht: Foris.

- Rizzi, Luigi. 1986. "Null Objects in Italian and the Theory of pro". *Linguistic Inquiry* 17.501-557.
- . 1997. "The Fine Structure of the Left Periphery". *Elements of Grammar* ed. by Liliane Haegeman, 282-337. Dordrecht: Kluwer.
- Roberge, Yves. 1991. "On the Recoverability of Null Objects". *New Analyses in Romance Linguistics* ed. by Dieter Wanner & Douglas A. Kibbee, 299-312. Amsterdam & Philadelphia: John Benjamins.
- . 2003. "Transitivité verbale, grammaticalisation et sémantisation". Paper presented at the Représentations du Sens Linguistique II colloquium, Montreal, May 2003.
- Troberg, Michelle. 2004. "La proposition topique-commentaire en moyen français: la reprise facultative du clitique accusatif". Ms., University of Toronto.
- Tuller, Laurice. 2000. "Null Objects in 'Deaf French'". Ms., Université de Tours.

# MICRO-PARAMETRIC VARIATION AND NEGATIVE CONCORD\*

VIVIANE DÉPREZ & FRANCE MARTINEAU  
*ISC, CNRS/Rutgers & Ottawa University*

## 0. Introduction

The purpose of this chapter is to compare the properties of negative concord (NC) in various French-related dialects to analyze better the factors that condition and limit observed micro-parametric variations. We examine both synchronic and diachronic Standard French (SF) in relation to Quebec French (QF) and French-based Creoles (FBC, e.g., Haitian Creole (HC)). The French dialects in (1) make use of apparently identical negative expressions (N-words, Laka 1990), but manifest strikingly different NC properties:

- (1) a. *Je (n') ai vu personne.* SF (synchronic and diachronic)  
b. *J'ai pas vu personne.* QF  
c. *Mwen pa wè pèsonn.* HC

For instance, the copresence of sentential negation with N-words is obligatory in FBC, fairly standard in QF, possible in diachronic SF, but excluded in contemporary SF. Furthermore, there is comparable diversity for other properties of NC, such as double negation, licensing in NPI (negative polarity item) contexts, locality, and modification possibilities. Following Jespersen (1917), it has been repeatedly claimed in the literature that observed cross-linguistic differences in NC could be attributed to the differing properties of sentential negation (Zanuttini 1997). In contrast, the purpose of this chapter is to demonstrate, following Déprez (2000), that the key to NC variation resides in the syntactic/semantic properties of the concord expressions themselves, that is, the N-words, and more specifically in their internal structure. The central theses that this work supports are as follows:

- The properties of concord are determined by the properties of N-words, not by the properties of sentential negation;

---

\* This research has been partly funded by a SHHRC grant (No. 410-2001-0119). Authors' names are listed alphabetically. We would like to thank LSRL participants for comments and suggestions. Many thanks also to our research assistant Mélanie Lamarche for help collecting the data.

- There is a syntax-semantic interface internal to N-expressions ; and
- Micro-parametric variation is governed by this interface.

Evidence for these points comes from a variety of converging sources. Our empirical comparison reveals that distinctions in the semantic properties of N-words go along with distinctions in their internal syntax. Déprez (2000) argued that SF N-words manifest characteristic properties of determiners and are located in the upper layers of DP structure and that, in contrast, HC N-words manifest properties that most resemble those of bare NPs and are located in the lowest layer of nominal structure. This chapter provides further confirmation for such a syntactic distinction among N-words based on diachronic and synchronic studies. As is well known, French N-words started out as indefinite positive expressions that gradually acquired a negative value. Our careful study of a number of French diachronic corpora<sup>1</sup> demonstrates that this change goes hand in hand with changes in syntactic properties. This study provides evidence that French N-words have undergone a syntactic change that affected their positions in the hierarchical DP structure and argue that this evolution, rather than the changes affecting sentential negation, is at the basis of the changing semantic properties of French NC. Similarly, QF NC is shown to manifest properties that are intermediate between those of SF and those of FBC, providing a synchronic instantiation of what was presumably a diachronic stage of SF. The synchronic and diachronic comparative evidence provided here questions traditional views of NC as being governed by the nature of sentential negation (Jespersen 1917; Zanuttini 1997) or as a mere lexical ambiguity between positive and negative N-words (Longobardi 1991; Herburger 2002). It shows instead that the key to NC variation is the internal syntax of N-words.

The chapter is structured as follows. First, arguments supporting the central thesis of Déprez (2000) are summarized, based on an empirical comparison of two extremely diverging cases of NC, SF and Martinique Creole. Although using virtually identical N-words, the two languages have strikingly distinct NC properties. Following Déprez (1999, 2000), these NC properties can be explained by the distinct structural nature of their N-words. Second, the

---

<sup>1</sup> Our diachronic data come from the following corpora: (a) Old and Middle French: *Textes de français ancien*, ARTFL Database; (b) 16<sup>th</sup>-century French: ARTFL Database; (c) 17<sup>th</sup>- and 18<sup>th</sup>- century French: ARTFL Database, Anthony Lodge's computerized corpus of 17<sup>th</sup>- and 18<sup>th</sup>- century texts, which he made available to us, Martineau's *Corpus de français familier classique*, and Martineau (to appear); (d) 19<sup>th</sup>- and 20<sup>th</sup>- century French: ARTFL Database and *Témiscouata* corpus.

proposed structural distinction is further put to the test with the study of Modern QF NC, which presents mixed properties with respect to SF and Martinique Creole. The final test offered is that of time, with a study of the diachronic evolution of *aucun*.

## 1. *Spelling out the background: Two extreme cases*

### 1.1 *N-words in SF*

This section reviews the properties of French N-words that form the core of our comparison.<sup>2</sup> Perhaps the most salient property of expressions like *rien* and *personne* is their apparent ambiguity. They seem, on the one hand, to have an intrinsic negative value (2a) and, on the other hand, to behave like positive dependent elements, that is, indefinite expressions in the scope of negation (2b):

- (2) a. *Qui as-tu rencontré ici? Personne.*  
 “Who did you meet here? No one.”  
 b. *Je n’ai jamais rencontré personne ici.*  
 “I have never met anyone here.”

In their indefinite value, N-words closely resemble NPI expressions like *qui que ce soit* “whoever” and *le moindre* “the least,” a similarity at the basis of their common analysis in much of the recent literature on NC. Many authors (Ladusaw 1992; Giannakidou 1998, among others) analyze NC as a special case of NPI licensing. These analyses, however, have neglected important differences that clearly distinguish these two types of expressions in SF. Example (3) shows that French N-words strongly contrast with NPI in being largely incompatible with sentential negation. More precisely, when co-occurring with sentential negation, French N-words have only a double-negation interpretation, not a concord one, so that example (3), as logic dictates, only has a positive interpretation:

- (3) *Il n’a pas rencontré personne.*  
 “He did not meet no one.”

Example (3) clearly shows that SF N-words, in contrast with NPIs and Italian/Spanish N-words, cannot take scope under sentential negation. Sequences of N-words like those in (4a) also clearly differ from sequences

---

<sup>2</sup> Deletion of *ne* in Modern French depends on social contexts. In SF, *ne* is still retained while in vernacular French, it is often deleted (Ashby 2001).

involving an NPI. In addition to a concord reading, paraphrased in (4b), (4a) can have a double negative reading, paraphrased in (4c):

- (4) a. *Personne ne commet aucune erreur.*  
 “No one makes no error.”  
 b. *Personne ne commet la moindre erreur.*  
 “No one makes any error/ the slightest error.”  
 c. *Tout le monde commet au moins une erreur.*  
 “Every one makes at least one error.”

A double-negative reading is never available for sequences containing an NPI (4b). There are yet further differences. For instance, (5a) shows that N-words can be modified by adverbs like *absolument* and *presque*, frequently (but wrongly) assumed to modify only universal quantifiers. Example (5b) shows that French NPIs, in contrast, do not support modifications of this kind:

- (5) a. *Je n’ai vu absolument/presque personne.*  
 “I have seen absolutely/almost no one.”  
 b. \**Je n’ai pas vu absolument/presque qui que ce soit.*  
 “I have not seen absolutely/almost anyone.”

In standard NPI licensing contexts, while NPIs have a positive indefinite interpretation, SF N-words maintain a context-independent negative interpretation, as shown in (6).

- (6) *Si tu vois qui que ce soit/ personne, dis-le-moi.*  
 “If you see anyone/ no one, tell me.”

Example (7) shows that the locality conditions governing concord readings in N-word sequences are far more restricted than those governing the dependency between an NPI and its licenser. The former are limited to a single proposition, whereas the latter span over a wider context, typical of long-distance dependencies. That is, as (7a) shows, the relationship [*ne* ... N-word] and concord readings must be established within the boundary of a single clause. For NPI, however, if c-command is respected, NPIs and licensers may occur in distinct clausal domains.

- (7) a. \**Je ne veux que tu fasses rien.*  
 Lit: “I (*ne*) want that you do nothing.”  
 b. *Je ne veux pas que tu fasses quoi que ce soit.*  
 Lit.: “I don’t want that you do anything.”  
 c. *Personne ne croit que tu as rien fait.* (only double-negative reading)  
 “No one thinks that you have done nothing.”

- d. *Personne ne croit que tu as fait quoi que ce soit.*  
 “No one thinks that you have done anything.”

Déprez (1997, 1999, 2000) proposed an analysis of NC that accounts for the properties in (2)-(7) and, specifically, for the double-negative readings that cannot be explained through NPI-type analysis. SF N-words are analyzed as cardinal negative quantifiers and argued to be similar in their properties to numerals like zero: Both have intrinsic quantificational and anti-additive properties. This analysis covers (3)-(6) immediately. Since N-words are intrinsically anti-additive (negative) (i.e., similar to standard English negative quantifiers *nothing*, *nobody*), double-negative readings are expected for both sequences of N-words and N-words co-occurring with sentential negation. Surprising from this point of view is the concord interpretation of (2b). To account for this reading, Déprez (1997, 1999, 2000) extends May's (1989) analysis of resumptive quantification. For May, any sequence of quantifiers [ $Q_1 \dots Q_n$ ] of similar nature can have two interpretations: (a) a sequential interpretation, in which quantifiers have scope over one another [ $Q_1 > \dots Q_n$ ], and (b) a resumptive interpretation obtained through the formation of a single polyadic quantifier that binds several variable at once [ $Q_{1\dots n} (x_1\dots x_n)$ ]. For negative quantifiers, the two interpretations derive two very distinct readings. The sequential interpretation derives a multiple-negation reading since the negative value of each quantifier is computed independently. The resumptive interpretation, in contrast, derives a concord reading because, in a single polyadic quantifier, the negative value is computed only once. Déprez' (1997, 1999, 2000) resumptive quantification analysis of French NC has a number of advantages over an NPI analysis. As May argues, only quantifiers sufficiently similar in nature can form a complex polyadic quantifier. Since sentential negation and N-words are clearly distinct elements—the former is a propositional operator, the latter a variable-binding quantifier—their association fails to support the formation of a polyadic quantifier. This correctly predicts that a sequence containing negation will always have a double-negative (sequential) interpretation, never a concord (i.e., polyadic) interpretation. The analysis further predicts that, if some operation could impose a scope structure on an N-word sequence, the double-negative reading would prevail, the resumptive reading being by definition scopeless with respect to a sequence. This prediction is verified. As Corblin (1994), Déprez (1997, 2000), and Vinet (1998) have observed, various stress/focus conditions on N-word sequences clearly favor a double-negation reading. Because of May's 'similarity condition,' the analysis also predicts that the more similar the members of a sequence, the easier it should be to form a polyadic quantifier. In



other words, in a sequence of N-words, similarity favors the concord reading and differences favor double negation. In conformity with this prediction, Déprez (2000) observes that in SF, the concord reading is most salient when participating N-words are all bare quantifiers (8a), and clearly harder in less parallel cases (8b):

- (8) a. *Personne n'a rien dit à personne.*  
 "No one said anything to anyone."  
 b. *Personne ne commet aucune erreur.*  
 "No one makes any/no mistake."

Finally, under the resumptive analysis, locality conditions are reduced to scope constraints. It is proposed that 'resumption' only succeeds among quantifiers with the same scope domain. Only quantifiers that are members of the same proposition can form a single polyadic quantifier, which correctly predicts that concord readings are bounded.

### 1.2 *N-words in Martinique Creole*

This section discusses the properties of N-words in Martinique Creole (MC) in comparison with those of SF. Quite strikingly, the MC N-words *personn*, *anyen*, and so on look, so to speak, lexically identical to those of SF and share the ability to have a negative value in isolation as an elliptical answer to a question (9). Important differences emerge, however, with N-words in sequences. As (10) shows, sequences of an N-word with the negation *pa* only have a concord reading, never a double-negation one. In declarative sentences, *pa* must in fact be present with N-words in all syntactic positions, as in (11). Clearly, MC N-words manifest a strong dependency on negation, just like NPI.

- (9) *Kimoun ou wè an fet tala? Personn.*  
 "Who did you see at this party? No one."  
 (10) *Man pa wè anyen.*  
 "I did not see anything."  
 (11) *Personn \*(pa) wè anyen.*  
 "No one (not) saw nothing."

Similarly, sequences of multiple N-words only have a concord reading in MC, never a double-negation reading. Furthermore, in typical NPI-licensing contexts like (12), MC N-words have a positive indefinite meaning, not a negative one.

- (12) *Es u we personn bon matin an?* (yes/no question)  
 "Did you see anyone this morning?"

Together, these facts suggest that MC N-words do not have an intrinsic negative meaning, but rather a context-dependent one. They always scope under negation and other licensers, appearing much closer to standard NPIs than to true negative quantifiers. Also similar to NPIs, MC N-words do not support adverbial modification by *almost* or *absolutely*, and they can be licensed in nonlocal contexts by a distant negation. Table 1 summarizes their properties as compared to SF.<sup>3</sup>

Properties	SF	Martinique Creole
Negative value	Yes	Yes
Double negation	Yes	No
Modification	Yes	No
Locality	Yes	No
NPI contexts	No	Yes
Negation compatibility	No	Yes

Table 1: *SF and Martinique Creole properties*

Table 1 shows that, apart from the negative value of (9), MC N-words and SF N-words have opposite properties. For the largely identical case of HC N-words, Déprez (1999) argues that the different NC properties stem from the distinct internal structure of the nominal expressions in these languages. Bare nouns without determiners are excluded in French argument positions (*\*J'ai mangé pommes* "I ate apples") but are extremely common in FBC. Although N-words such as *rien* and *personne* seemingly take the form of bare nouns in both languages, Déprez shows that they have in fact opposite internal structures. French N-words are determiner-like in nature and occupy a high position in the functional structure of nominal expressions (13a). Creole N-words behave more like bare nouns, with a low position in the nominal structure and a null determiner (13b):

- (13) a. [<sub>DP</sub> *personne* [<sub>NP</sub> Ø ]] French N-words  
 b. [<sub>DP</sub> Ø [<sub>NP</sub> *personn* ]] MC N-words<sup>4</sup>

French N-words are autonomous quantificational determiners or bare quantifiers, while Creole N-words are context-dependent expressions for their licensing and interpretation, the null determiner being a variable that requires appropriate binding. Accordingly, there are two distinct types of NC: (a) a

<sup>3</sup> For a discussion of comparable facts in HC, see Déprez (1992, 1999). See also DeGraff (1993) for a different analysis of NC in HC.

<sup>4</sup> The DP structure in these schemas is simplified for expository reasons. In general, D positions in this chapter should be understood as referring to various functional non-N positions, such as NumP.

quantificational type that arises between autonomous negative quantifiers when a single polyadic quantifier is formed and (b) a dependent type that spells out a variable-binding relation semantically close to that of NPI licensing. These two types of negative concord illustrate the two extreme structures that N-words may take and form the two extreme poles of negative relations, but variation between the two poles may occur as determined by the inner syntax of N-words and a corresponding changing semantic interpretation. The higher an N-word is in the functional DP structure, the closer it will be to negative quantifiers, and vice versa. Within this perspective, a change from a positive interpretation (MC) to a negative one (SF) can be understood as a change in the internal structure of N-words that arises from the movement of the N-word up the functional DP structure.

## 2. *Extending the results*

### 2.1 *QF N-words: An intermediate case*

N-words in QF are lexically identical to those of SF. Moreover, like them, they can have a negative value in both interrogative and declarative contexts, as in (14):

- (14) a. *As-tu vu quelqu'un ici hier soir? Personne.*  
 "Did you see anyone here yesterday night?"  
 b. *Il a aucun droit.*  
 "He has no rights."

Contrasted with SF, however, QF N-words are compatible with the sentential negation *pas*, as in (15).<sup>5</sup>

- (15) a. *Je ne peux pas avoir rien.* (Témiscouata)  
 "I cannot have anything."  
 b. *Je n'n ai pas aucune ouvrage payante.* (Témiscouata)  
 "I don't have any paying work."

The copresence of *pas* with a QF N-word induces a concord reading, not a double-negative one. Yet the presence of the negation is optional in QF, not obligatory. There is an interesting asymmetry in the distribution of negation, similar to the distributional asymmetry observed in Italian (Zanuttini 1997). *Pas* can co-occur with an N-word in postverbal position (16), but not with one that precedes the verb (16) (see also Di Sciullo & Tremblay 1996).

- (16) \**Rien (ne) m'arrive pas.*  
 "Nothing happens to me."

<sup>5</sup> In Modern QF, deletion of *ne* is rather systematic (Sankoff & Vincent 1977).

In QF, sequences of N-words like the ones found in example (17) can optionally include the sentential negation *pas*, but do not generally allow for a double-negative reading (17b):

- (17) a. *Personne (n') a (pas) pu rien nous dire.*  
 “No one could tell us anything.”  
 b. *Aucun enfant (n') a (pas) rien mangé.* (no double negation)  
 “No child has eaten anything.”

Moreover, in characteristic NPI contexts, N-words can still have a positive reading in QF, as in (18):

- (18) a. Yes/no question:  
*T'as-tu vu aucun chien dans les parages?*  
 “Did you see any dog in the neighborhood?”  
 b. Conditional:  
*Si tu vois aucun étudiant, appelle-nous.*  
 “If you see any student, call us.”  
 c. Negative predicates:  
*Elle refuse de dire aucun mot à la police.*  
 “She refuses to say anything to the police.”  
 d. Temporal adverbs:  
*Avant de faire aucune chose de final, tu ferais bien mieux de réfléchir.*  
 “Before doing anything final, you’d better think about it.”

Regarding modification possibilities and locality constraints, QF seems to present a mixed picture. When N-words occur alone, without *pas*, they seem to have an intrinsic negative value and, accordingly, can be modified by *almost* or *absolutely* (19a). This modification, however, is unacceptable in the copresence of an N-word and the sentential negation *pas* (19b):

- (19) a. *Je (n') ai rencontré presque aucun chum.*  
 “I met almost no friend.”  
 b. \**Je (n') ai pas rencontré presque aucun chum.*  
 “I did not meet almost any friend.”

Similarly for locality conditions, the dependent N-word can occur a long distance from *pas* if it is present, that is, in a lower proposition (20a). However, if only *ne* is present, the distance from a related N-word is restricted to a single propositional domain (20b).

- (20) a. *Il (ne) faut pas [que je prenne aucun coup ce soir].*  
 “It must not be that I take any hit tonight.”  
 b. \**Il ne faut [que je prenne aucun coup ce soir].*

These facts suggest that QF N-words are ambiguous. In some cases, they function like independent negative quantifiers, as in the simple declarative context in (14b). In other cases, as in NPI contexts, they function as dependent elements similar to standard NPI. QF N-words thus seem to have properties that are intermediate between those of SF and those of MC. Table 2 highlights this situation.

Properties	SF	MC	QF
Negative value	Yes	Yes	Yes
Double negation	Yes	No	No
Modification	Yes	No	Yes/No
Locality	Yes	No	Yes/No
NPI contexts	No	Yes	Yes
Sentential negation compatibility	No	Yes	Yes but not obligatory

Table 2: *SF, Martinique Creole, and QF properties*

The distribution in Table 2 would follow if QF N-words had an ambiguous structure, occurring either in a determiner-like position as in (13a) or dominated by a null determiner as in (13b). This structural ambiguity is expected if, as hypothesized by Déprez (2000), N-words can move up the DP structure. Movement to D would be obligatory in SF, impossible in MC, and optional in QF, deriving the observed differences. From this point of view, the existence of an intermediate case like QF, where an internal movement of N-words is possible but not obligatory, brings further evidence to the proposed structural distinction. However, recent syntactic models have suggested that movement is never optional. A closer look shows that the QF NC data may in fact confirm this view. Note that if QF N-words could freely have either structure (13a) or (13b), they should have the properties of SF and MC combined. But this is not exactly the case. As noted earlier, sequences of N-words in QF do not seem to allow a double-negation reading. This apparent anomaly suggests that structure (13b) is obligatory whenever an N-word is c-commanded by another negative element, so that the distribution of (13a) and (13b) is governed by the principle in (21):

- (21) When null D is licensed, N-word movement to D does not occur, otherwise it is obligatory.

In other words, N-word movement is a last-resort strategy that eliminates (13b) wherever null D fails to be licensed. On this view, movement is obligatory when it is necessary, and impossible otherwise, as expected from a Minimalist Program perspective. This approach makes an interesting prediction. Within a sequence, an N-word c-commanding another should have

the structure (13a), and the second one the structure (13b). Recall the aforementioned asymmetry in the co-occurrence of QF N-words with sentential negation. This asymmetry can now be explained in the following way: An N-word in a subject position c-commands negation but is not c-commanded by it, and therefore cannot be licensed by it. Such an N-word must then have the structure in (13a), which predicts that it will function like a SF N-word, expectedly disallowing the copresence of negation. On the other hand, as N-words in complement position are c-commanded by negation, they can have the structure (13b). Their null D is licensed so that co-occurrence with negation is expected. (The reader is invited to verify that the proposal accounts for the remaining observed properties of QF negative concord, as space limitations do not permit us to do so here.)

## 2.2 *Time will tell*

Perhaps the best-known feature of the evolution of French N-words is their semantic change from positive to negative terms. *Pas* started out as the positive noun “step,” *rien* as “thing,” *personne* as “people,” and so on, but they all ended up negative, meaning respectively “no,” “nothing,” and “no one.” While carefully retracing the diachrony of these terms, we have observed that this well-known semantic change goes hand in hand with much less discussed morphosyntactic changes. Since space constraints prevent a detailed exposition, this chapter presents a main outline of this co-evolution, focusing more particularly on the N-word *aucun* (see Martineau & Déprez to appear, for *rien*). In unpublished work, Schnedecker and Prévost (2002) also studied the evolution of *aucun*, focusing solely on its morphosyntactic properties. Their thorough data complemented ours.<sup>6</sup> The main empirical results of our study and the analysis suggested for these results are respectively displayed in Tables 3 and 4.

OF-MF	1. <i>aucun</i> → Sing/Pl & Mod/Pro	Positive
16 <sup>th</sup> , 17 <sup>th</sup> , 18 <sup>th</sup> c.	2. <i>aucun</i> → Decrease in Pl & Pro	Indeterminate
19 <sup>th</sup> -20 <sup>th</sup> c.	3. <i>aucun</i> → Reduction to Sing Det	Negative

Table 3: *Main results; the change in semantic value correlates with a morphosyntactic change in N-words*

<i>aucun</i> : D <sup>0</sup> /Spec DP +Q	Positive	Positive polarity item
<i>aucun</i> : Adjectival - Q		Negative polarity item
<i>aucun</i> : D <sup>0</sup> +Q	Negative	Negative quantifier

Table 4: *Suggested analysis, syntactic and semantic changes*

<sup>6</sup> If not specifically mentioned, the data are from the corpora listed in fn. 1; otherwise, we indicated Schnedecker and Prévost (2002). Sometimes, as we indicated, we present their data following our own format (e.g., Table 6).

Table 3 shows that *aucun* had a positive meaning in Old French (OF) and Middle French (MF) that seems linked to an increased use of its plural and pronominal forms, adding up to almost half of its total uses by the MF period. That is, *aucun* in MF occurred almost as often as a pronoun as it did as a noun modifier and as much in the plural as in the singular. In Table 4, *aucun*, in this initial phase, occupied a determiner position in the nominal structure, and is a positive independent existential pronominal/determiner quantifier, similar to the current *quelqu'un* “someone,” itself a positive polarity item. In a subsequent phase, we observe that the plural and the pronominal forms start to decrease, while the meaning of *aucun* changes to become indeterminate and context dependent (NPI). *Aucun* becomes a dependent element with no intrinsic quantificational force. During this phase, *aucun* occupies a low adjectival position in the nominal structure, leaving the D<sup>0</sup> position empty. Characteristically during this period, constructions like (22) are common where *aucun* is used in opposition to *autre* (see Schnedecker & Prévost for the data), analyzed as an adjectival element in Eguren and Sanchez (this volume):

(22) *Les aucuns... les autres...*

In its final phase, *aucun* is almost exclusively singular and a determiner, having essentially lost the capacity to be pluralized and be used as a pronoun. *Aucun* becomes independent again and negative on its own, taking on the intrinsic quantificational properties that characterize current French N-words. In this last phase, *aucun* climbed back up the functional structure of its nominal constituent, becoming invariable and strongly quantificational, similar morphosyntactically to the universal quantifier *chacun*. The evolution of *aucun* corresponds to a cycle during which it changes position inside the nominal structure. The study proposes that *aucun* begins its life as a positive specifier of DP with an intrinsic quantificational force and shows that *aucun* at this time could take scope over negation, like a typical positive polarity item. Later on, *aucun* moves down the nominal structure, taking on the value of an adjective without quantificational force. This downward move correlates with the decrease of its pronominal use. Adjectival *aucun* plausibly has a null determiner as in (13b), which accounts for its context-dependent interpretation, positive in NPI contexts, negative under the scope of negation. Finally, *aucun* moves back up the DP, this time perhaps to a D/Num head position, taking on a negative meaning that seems linked in part to its incapacity to be pluralized, while having a distributive meaning and full quantificational force. This move back up the nominal structure could be prompted by the gradual loss of null determiners in French.

This is obviously a schematized and simplified approach to the evolution of *aucun*. Often, the different stages of this evolution are not as clear as the analysis predicts. However, this proposal has the merit of linking syntactic changes with semantic changes in a way that accounts for differences observed elsewhere and is compatible with the diachronic facts, to which we now turn.

In OF and MF, *aucun* is positive with a meaning comparable to that of *some*, and is rarely used in polarity or negative contexts.

Periods	Positive	Polarity	Negative
OF and MF	73.5% (463/630)	15.9% (100/630)	10.6% (67/630)

Table 5: *Frequency of aucun in positive, polarity, and negative contexts in OF and MF*

*Aucun* is mostly used without *ne* (23a) and, when it occurs in a negative context it can maintain a positive meaning. Notably, in (23b) *aucun* has scope over negation, behaving like a positive polarity item.

- (23) a. *Mais aucuns quis vit esbuschier*  
*Le curut al rei acuinte (Brut, Buridant 2000:179)*  
 “But some who saw them hide ran to announce it to the king”
- b. *Or ne porront pas dire aucuns ke j'ai antés*  
*Ke d'aler a Paris soie por nient vantés (Feuillée, Buridant 2000:179)*  
 “Some of those I have known could not say that I boasted in vain of having gone to Paris.”

In OF and MF *aucun* can be a determiner or a pronoun and can be singular or plural. While the singular determiner form dominates in OF, there is a gradual increase of the pronominal and plural forms so that in MF, the distribution of all these distinct forms becomes even (Table 6).

Periods	Sing	Pl	Det	Pro
OF	88% (899)	12% (129)	61.5% (632)	38.5% (396)
MF	54% (5085)	46% (4331)	50.1% (2595)	49.9% (2579)

Table 6: *Frequency of aucun as a singular/plural term and as a determiner/pronoun in OF and MF (adapted from Schnedecker & Prévost 2002)*

*Aucun*, be it a noun modifier (24a) or a pronoun (24b), occasionally appears with a preceding determiner.

- (24) a. *Premièrement, les aucuns espreviers se perchent tout droit et sont moult esveilliez*  
*(Le Mesnagier de Paris 1394, from Schnedecker & Prévost 2002)*  
 “First, some sparrow hawks perch straight up and are very awake.”
- b. *Nus ne se doit merveiller se les aucunes se departent de leur maris (Coutumes de Beauvaisis, Moignet 1984:176)*  
 “No one must be surprised if some get rid of their husbands.”



Occasionally, *aucun* occurs postnominally as a noun modifier. This post-nominal construction undergoes a sharp increase during the 16<sup>th</sup> century (see Table 11), a fact we believe to be significant, and which we return to later.

- (25) *Mais il n'y a occasion aucune entre nous deux* (Navarre, *La Coche*:186, ARTFL)  
 “But there isn’t occasion any between both of us.”

It is clear that in constructions such as (24a) and (24b), *aucun* does not occupy a determiner position, the D position being respectively occupied by an overt determiner in (24a) and presumably empty in (25). *Aucun* must then occupy a lower position in the DP structure as a modifier of some sort. The increasing frequency of such constructions marks the start of a structural change in the position of *aucun* within DP, beginning a descent from a positive SpecD to a lower adjectival-like position.

In correlation with this structural change, the 16<sup>th</sup> and 17<sup>th</sup> centuries manifest a change in the interpretation of *aucun*. There is a steady increase in the indeterminate interpretation, as *aucun* now appears more and more frequently in negative and polarity contexts and much less as a positive element. Table 7 provides an interesting contrast with Table 5. Notice that, from the 16<sup>th</sup> century, the positive interpretation of *aucun* is no longer primary, and that by the 17<sup>th</sup> century, it has dwindled to an almost insignificant proportion.

Periods	Positive	Polarity	Negative
16 <sup>th</sup> c.	21.3 % (25/117)	30% (35/117)	48.7% (57/117)
17 <sup>th</sup> c.	3.5% (7/200)	27.5% (55/200)	69% (138/200)

Table 7: Frequency of *aucun* in positive, polarity, and negative contexts, 16<sup>th</sup> and 17<sup>th</sup> c.

Some of the most commonly found polarity contexts are illustrated in (26) (all examples from ARTFL; cf. also Fournier 1998):

- (26) a. Yes-no question:  
*Penses-tu qu'aucun d'eux songe à nous faire mal?*  
 b. Conditional:  
*s'il est aucun respect ni pouvoir qui m'arrête...*  
 c. Factive predicates:  
*J'aurai regret d'en épargner aucune*  
 d. Negative predicates:  
*Vous n'avez pas lieu d'en prendre aucun soupçon*  
*Dieu ne vous a pas mise en ce monde pour aucun besoin*  
 e. Consecutive:  
*Ma fille est d'une race trop pleine de vertus pour se porter jamais à faire aucune chose*

The 16<sup>th</sup> and 17<sup>th</sup> centuries are characterized by a growing systemization of the correlation between the morphosyntactic forms of *aucun* with its interpretations. On the morphosyntactic front, Table 8 shows that, while the proportion of singular forms of *aucun* steadily increases, the plural forms begin to recede. Simultaneously, the noun-modifier use increases and the pronominal use decreases.

Periods	Sing	Pl	Mod	Pro
16 <sup>th</sup> c.	71.8% (84)	28.2% (33)	65% (76)	35% (41)
17 <sup>th</sup> c.	93.5% (187)	6.5% (13)	86% (172)	14% (28)

Table 8: *Frequency of aucun as singular/plural term, noun modifier/pronoun, 16<sup>th</sup> and 17<sup>th</sup> c.*

If pronouns are in D<sup>o</sup>, the observed decrease of this form supports the hypothesis that *aucun* is slowly moving away from a D position to take on a modifier role within DP. It is then particularly interesting to note that the changes in the interpretation of *aucun* observed in Table 7 correlate with the changes of forms observed in Table 8. Table 9 shows that the increasing singular noun-modifier form is more frequently used in negative and polarity contexts, and very rarely in positive contexts.

Periods		Positive	Polarity	Negative
16 <sup>th</sup> c.	Sing	0% (--)	21.1% (16)	60.5% (46)
	Pl	10.5% (8)	5.3% (4)	2.6% (2)
17 <sup>th</sup> c.	Sing	0% (--)	28.5% (49)	65.7% (113)
	Pl	2.3% (4)	1.2% (2)	2.3% (4)

Table 9: *Frequency of aucun as a noun-modifying form in positive, polarity, and negative contexts, 16<sup>th</sup> and 17<sup>th</sup> c.*

Table 10 shows that the decreasing pronouns are more frequently plural in positive readings, and singular in polarity and negative contexts. A correlation is clearly emerging between the noun-modifier and singular uses of *aucun* and its negative and polarity readings. Positive readings in contrast become confined to the decreasing plural and pronominal forms.

Periods		Positive	Polarity	Negative
16 <sup>th</sup> c.	Sing	2.5% (1)	34.2% (14)	17% (7)
	Pl	39% (16)	2.4% (1)	4.9% (2)
17 <sup>th</sup> c.	Sing	0% (--)	14.3% (4)	75% (21)
	Pl	10.7% (3)	0% (--)	0% (--)

Table 10: *Frequency of aucun as pronoun in positive, polarity, and negative contexts, 16<sup>th</sup> and 17<sup>th</sup> c.*

- (27) *Aucuns disoient que c'etoient des pretres* (La Fontaine, *Fables*, Fournier 1998:229)  
 “Some say that these were priests.”

Further interesting evidence of a correlation between morphosyntax and interpretation emerges when the noun-modifier *aucun* is focused on. As noted earlier, there is a significant increase in the postnominal position of *aucun* in the 16<sup>th</sup> century. Even more interesting is the fact that, as Table 11 shows, this postnominal position occurs more frequently in a negative or polarity context than in a positive one.

16 <sup>th</sup> c.	Positive	Polarity	Negative
<i>aucun</i> N	8	9	24
N <i>aucun</i>	--	11	24

Table 11: Raw numbers of *aucun* in pre-and postnominal positions in 16<sup>th</sup> c.

In positive contexts, the prenominal position is clearly favored, while in negative and polarity contexts, slightly more than half of the occurrences of *aucun* are in postnominal position. This tendency to correlate the postnominal position with the polarity/negative interpretation strongly supports the idea that *aucun* acquires an indeterminate context-dependent interpretation when it is adjectival, the postnominal position being, indeed, unambiguously adjectival. With adjectival *aucun*, DP plausibly contains a null determiner, as in the structure (13b), which predicts a context-dependent interpretation, if null determiners are binding-requiring variables.

The 19<sup>th</sup> and 20<sup>th</sup> centuries are most clearly characterized by the disappearance of both the plural and the pronominal forms. In other words, the singular noun-modifier use that started to increase in the 16<sup>th</sup> century now becomes fully dominant (cf. Table 12).

Periods	Sing	Pl	Det/Adj (masc)	Pro (masc)
19 <sup>th</sup> -20 <sup>th</sup> c.	99.9% (23,574)	0.1% (189)	84% (1,274)	16% (244)

Table 12: Frequency of *aucun* as a singular/plural term; as a determiner/adjective/pronoun in 19<sup>th</sup> and 20<sup>th</sup> c. (adapted from Schnedecker & Prévost 2002)

The plural use has almost entirely disappeared. Interestingly, 20<sup>th</sup>-century *aucun* is not strictly singular because it occurs with mass nouns: It appears underspecified for number, as in (28):

- (28) *Il n'a aucun argent.*  
 “He has no money.”

Table 13 shows (compare with Table 7) that the positive uses of *aucun* have all but disappeared. The corpora we consulted (cf. fn. 1) contain no occurrence of it, yet for most current speakers a late-developing and particularly interesting positive form of *aucun* is still available, and it must be preceded by the marker *d'*, as in (29).

- (29) *D'aucuns pensent que l'on devrait toujours éviter la guerre.*  
 "Some think that war should always be avoided."

Periods	Positive	Polarity	Negative
18 <sup>th</sup> c.	0%	20.4% (98)	79.6% (479)
19 <sup>th</sup> c.	0%	17.1% (81)	82.9% (390)
20 <sup>th</sup> c.	0%	15.8% (38)	84.2% (201)

Table 13: Frequency of *aucun* in positive, polarity, and negative contexts, 18<sup>th</sup>, 19<sup>th</sup>, and 20<sup>th</sup> c.

The form in example (29) is quite interesting as it shows that the only modern positive form of *aucun* is one in which the determiner position is occupied.<sup>7</sup> *Aucun* may be in a low DP position but, as D<sup>o</sup> is occupied, the interpretation is fixed, not context dependent.

Table 13 also shows that the use of *aucun* in polarity contexts is decreasing. A closer look at the data reveals that, for the 18<sup>th</sup>, 19<sup>th</sup>, and 20<sup>th</sup> centuries, the main polarity context used is with the preposition *sans*. When this context is excluded (it was analyzed as a possible resumptive structure in de Swart & Sag 2002), the frequency of *aucun* in polarity contexts nears 0% in the 20<sup>th</sup> century.

In the 17<sup>th</sup> century, *aucun* was still used as a polarity item, but in the 18<sup>th</sup> century this use decreased, while the use of purely negative meaning was increasing. As shown in Martineau and Mougeon (2003), deletion of *ne* became prevalent only during the 19<sup>th</sup> century. In other words, *ne* deletion seems to have followed the meaning change in *aucun*, not to have caused it. (See also Martineau & Vinet to appear for an analysis of presence/absence of *ne* in interrogative and exclamative contexts.) It thus seems that the rise of the negative meaning of *aucun* correlates not with the weakening of negation, but more significantly with the disappearance of its plural form and perhaps more generally with the disappearance of null determiners in the language. Its invariability suggests that *aucun* may have climbed back up the nominal structure, again becoming an independent quantifier similar to *chacun*. Its resulting negative value may perhaps be explained in terms of a proposal by

<sup>7</sup> There are also positive forms for N-words like *rien* and *personne* in SF all with an overt determiner: *Un petit rien le dérange* "A little something bothers him," *Une personne...* "one/a person."

Postal (2004), who suggests that all indefinite terms contain sets of negative features that can either annihilate one another (positive value), create a negative dependency (NPI), or be reinforced (pure negative value). However, why the negative value of these features should correlate with a high position in the determiner structure still needs to be understood better, as does the role that number plays in these instances.<sup>8</sup> At this point in our diachronic inquiry, however, this question must be left for further study. To conclude this section, Table 14 summarizes the suggested analysis of the facts presented.

---

OF	<i>Aucun</i> = Positive existential Q = PPI
MF	<i>Aucun</i> = D <sup>0</sup> increase of pronoun (and plural)
16 <sup>th</sup> c.	<i>Aucun</i> becomes adjectival = weak or strong numeral indefinite, quantificational or predicative. When weak, it no longer is in D <sup>0</sup> (postnominal).
17 <sup>th</sup> c.	Adjectival <i>aucun</i> is increasing: Pronoun and plural are losing ground. D <sup>0</sup> is emptied. A dependency emerges that parallels the dependency of bare nouns.
18 <sup>th</sup> -19 <sup>th</sup> c.	Disappearance of <i>aucun</i> in D <sup>0</sup> . D <sup>0</sup> is null. It must be legitimized (NPI) or filled ( <i>d'aucuns</i> ).
20 <sup>th</sup> c.	New change : Invariable (singular) <i>aucun</i> climbs back into D <sup>0</sup> but now with a negative value.

---

Table 14: *Scenario of the evolution of aucun*

### 3. Conclusion

This chapter presented a comparative landscape of the NC properties in a variety of francophone dialects and historical states and has proposed an analysis that relies on the internal structure of N-words to account for the observed diversity. Two structures have been proposed: (13a) locates the N-word in a quantificational determiner-like position and (13b) assumes, on the contrary, that the N-word is dominated by a null determiner. Distinct semantic properties are associated with each structure and movement internal to the nominal structure of N-words could account for the variations observed synchronically and diachronically among the distinct dialects. A more detailed study of the syntactic properties of N-words would be useful to better support the suggested analysis, but we hope to have provided sufficient evidence to establish the value of our proposed approach.

---

<sup>8</sup> A full explanation as to how the negative value arises is yet to be provided. We could, as others do, simply add or activate an arbitrary [+Neg] feature to motivate the movement of N-words to D, but pending a valid semantic analysis of how negation incorporates to N-words, this would amount to a mere restatement of the facts.

## REFERENCES

- ARTFL Database. <http://humanities.uchicago.edu/orgs/ARTFL/> (January 2003-April 2004)
- Ashby, William. 2001. "Un nouveau regard sur la chute du *ne* en français parlé tourangeau: s'agit-il d'un changement en cours?". *Journal of French Language Studies* 11.1-22.
- Buridant, Claude. 2000. *Grammaire nouvelle de l'ancien français*. Paris: Sedes.
- Corblin, Francis. 1994. "Multiple Negation Processing". Technical report no. HCRC/RP-62, Human Communication Research Centre, University of Edinburgh.
- DeGraff, Michel. 1993. "A Riddle on Negation in Haitian". *Probos* 5.63-93.
- Déprez, Viviane. 1992. "Is Haitian Creole Really a Pro-Drop Language?". *Travaux de recherche sur le Créole Haïtien 11* ed. by Claire Lefèbvre, 23-40. Montreal: Université du Québec à Montréal.
- , 1997. "A Non Unified Analysis of Negative Concord". *The Syntax and Semantics of Negation* ed. by Danielle Forget, Paul Hirschbühler, France Martineau, & Maria-Luisa Rivero. Amsterdam & Philadelphia: John Benjamins.
- , 1999. "The Roots of Negative Concord in French and French Based Creoles". *Language Creation and Language Change: Creole, diachrony and development* ed. by Michel DeGraff, 375-428. Cambridge, MA: MIT Press.
- , 2000. "Parallel (A)symmetries and the Structure of Negative Expressions". *Natural Language and Linguistic Theory* 18.253-342.
- de Swart, Henriette & Ivan A. Sag. 2002. "Negation and Negative Concord in Romance". *Linguistics and Philosophy* 25.373-417.
- Di Sciullo, Anna-Maria & Mireille Tremblay. 1996. "Configurations et interprétation: les morphèmes de negation". *Recherches Linguistiques de Vincennes* 25.27-52.
- Fournier, Nathalie. 1998. *Grammaire du français classique*. Paris: Belin.
- Giannakidou, Anastasia. 1998. *Polarity Sensitivity as (Non)-Veridical Dependency*. Amsterdam & Philadelphia: John Benjamins.
- Herburger, Helena. 2002. "The Negative Concord Puzzle Revisited". *Natural Language Semantics* 9.289-233.
- Jespersen, Otto. 1917. *Negation in English and Other Languages*. Copenhagen: Host.
- Ladusaw, William. 1992. "Expressing Negation". *Ohio State University Working Papers in Linguistics* 40.237-259.
- Laka, Itziar. 1990. *Negation in Syntax: On the nature of functional categories and projections*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Longobardi, Pino. 1991. "Island Effects and Parasite Constructions". *Logical Structure and Linguistic Structure* ed. by James T. Huang & Robert May, 149-196. Dordrecht: Kluwer.
- Martineau, France. Corpus de Français Familier Classique (series of 500 informal letters written by low-educated people from Quebec, Acadie, and northwestern France). <http://www.uottawa.ca/academic/arts/lettres/nf> (January 2003-April 2004)
- , To appear. "Un corpus de textes français pour l'analyse de la variation diachronique et dialectale". *Lexicométra. L'analyse de données textuelles: de l'enquête aux corpus littéraires*.
- Martineau, France & Viviane Déprez. To appear. "Pas Aucun/Pas Rien en français classique: variation dialectale et historique". *Langue Française*.
- Martineau, France & Raymond Mougeon. 2003. "Sociolinguistic Research on the Origins of ne Deletion in European and Quebec French". *Language* 79.118-152.

- Martineau, France & Marie-Thérèse Vinet. To appear. "Negation in the Domain of the Left Periphery in Middle and Modern French". *Grammaticalization and Parametric Change* ed. by Montse Batllori & Francesc Roca. New York: Oxford University Press.
- May, Robert. 1989. "Interpreting Logical Form". *Linguistics and Philosophy* 12.387-437.
- Moignet, Gérard. 1984. *Grammaire de l'ancien français*. Paris: Klincksieck.
- Postal, Paul M. 2004. "A Remark on English Double Negatives". *Lexique, Syntaxe et Lexique-Grammaire/Syntax, Lexis & Lexicon-Grammar: Papers in honour of Maurice Gross* ed. by Christian Leclère, Eric Laporte, Mireille Piot, & Max Silberztein, 497-508. Amsterdam & Philadelphia: John Benjamins.
- Sankoff, Gillian & Diane Vincent. 1977. "L'emploi productif du ne dans le français parlé à Montréal". *Le Français Moderne* 45.243-254.
- Schnedecker, Catherine & Sophie Prévost. 2002. "Aucun(e)s/d'aucun(e)(s)/les aucun(e)s: Évolution du français médiéval au français moderne". Paper presented at the Séminaire Diachronie du Français, Paris, May 2002.
- Témiscouata Corpus (F. Dupuis, M. Faribault, & C. Portelance; series of informal letters written by low-educated people from Quebec at the beginning of the 20<sup>th</sup> century). <http://www.ling.uqam.ca/forum/corpus/messages/3.html> (January 2003-April 2004)
- Vinet, Marie-Thérèse. 1998. "Contrastive Focus, French N-words, and Variation." *Revue Canadienne de Linguistique* 43.121-141.
- Zanuttini, Rafaella. 1997. *Negation and Clausal Structure: A comparative study of romance languages*. New York: Oxford University Press.

# CONTRAST AND ADDITION IN ROMANCE

## A CASE STUDY IN MICROVARIATION\*

LUIS EGUREN & CRISTINA SÁNCHEZ

*Universidad Autónoma de Madrid & Universidad Complutense de Madrid*

### 0. *Introduction: Parameters and microparameters*

As is well known, principles and parameters theory aims to account for linguistic variation by means of a set of options (or parameters) left open by UG, which are considered to be confined to the lexicon. Within this general framework, a useful distinction has been made between large-scale, major parameters and fine-grained, minor patterns of variation or ‘microparameters’ (see, e.g., Baker 1996:496; Kayne 1996).<sup>1</sup> Major parameters and microparameters are alike in that both refer to formal properties of lexical items.<sup>2</sup> The difference between the two seems then to be one of degree or scope (see Holmberg & Sandström 1996). Major, standard parameters, like the *pro*-drop parameter: (a) have a complex cascade effect on the grammar of a language, usually affecting many constructions, (b) are (therefore) resistant to change, and (c) are (therefore) resistant to dialectal variation. Microparameters, on the other hand: (a) have a more limited impact on the grammar of a language, usually affecting a single lexical item, (b) are particularly vulnerable to change, and (c) are likely to show dialectal variation. In this chapter we present a case study on

---

\* We would like to thank our French, Italian, Catalan, Portuguese, and Romanian informants, Théophile Ambadiang, Lorenzo Bartoli, Filipa de Paula-Soares, Leo Luceri, Fina Llorca, Nadie Nerbesson, Marie-Lise Rebeyrol, Juan Ribera, Laurence Rouanne, Arlette Séré, and Laura Eugenia Tudoras. We are also very thankful to Olga Fernández Soriano and three anonymous reviewers for their comments. All remaining errors are ours. The research underlying this work has been partly supported by a grant to the project “La variación gramatical: variación micro y macroparamétrica en la morfología y la sintaxis. Teoría, descripción y aplicaciones” (BFF2000-1307-C03-02).

<sup>1</sup> As mentioned in the text, microparameters are here defined in relation to standard, major parameters. The term ‘microvariation’ can also jointly refer to both major and minor parameters, in contrast to macrovariation studies (Baker 1996).

<sup>2</sup> A common view within the principles and parameters model is that all linguistic variation is to be formulated in terms of formal properties of ‘functional’ categories. We believe this idea to be too restrictive, at least for microvariation studies. In this chapter we present a case study in microvariation that makes use of categorial feature changing ([+A] > [+D]/[+Deg]), and does not *stricto sensu* resort to formal properties of functional categories.



the expression of contrast and addition in Romance in which all three of these features of microvariation are illustrated and addressed. Due to space limitations, we will focus on Spanish and French, but our basic insights can be extended to other Romance languages as well.

Both Spanish and French have a prenominal ‘predicate of contrast’ (*otro*, *autre*) that makes the modified NP different from a given one belonging to the same set:<sup>3</sup>

- (1) a. {*El/Este*} *otro estudiante ha aprobado el examen.*  
 b. {*L/Cet*} *autre étudiant a passé son examen.*  
 “{The/This} other student has passed his exam.”

However, two major loci for microvariation between Spanish and French—and among Romance languages more generally—can be found with respect to this lexical item, both of them related to the (im)possibility of its becoming a functional category (i.e., either a determiner or a degree word). In sections 1 and 2, it will be shown that Spanish *otro* is both a prenominal predicate of contrast and a determiner, whereas French *autre* is an adjective and also behaves as an additive degree operator. In section 3, we will examine the French-like, adjectival, and degree properties *otro* had in Old Spanish, which are still present in some varieties of American Spanish today.

### 1. *Determiner otro, adjectival autre*

Its reference-related meaning (i.e., the fact that it denotes a referentially distinct individual with a shared description) may first make it possible for an adjectival predicate of contrast to have referential force on its own. This is the case for Spanish *otro*. As shown in (2) and (3), *otro* can be a determiner that licenses a nominal expression in an argument position, whereas French *autre* is never used that way:

- (2) a. *Otro estudiante ha aprobado el examen.*  
 b. \**Autre étudiant a passé son examen.*  
 “Another student has passed his exam.”  
 (3) a. *Otros estudiantes han aprobado el examen.*  
 b. \**Autres étudiants ont passé leur examen.*  
 “Other students have passed their exam.”

The paradigm in (2) and (3) strongly parallels those in (4)-(11), which illustrate how *otro* is deprived of some of the clear-cut adjectival properties

<sup>3</sup> A comprehensive study of Spanish *otro* can be found in Eguren and Sánchez (2003). In Van Peteghem (1997a, 1997b), French *autre* is approached from a discourse perspective.

displayed by *autre*. As can be seen in (4) and (5), *autre* can first be the input for the formation of a *-mente* adverb:

- (4) a. \**Lo quiero hacer otramente.*  
 b. *Je veux le faire autrement.*  
 “I want to do it in another way.”
- (5) a. \**Tienes que volver pronto; otramente, no cenarás.*  
 b. *Tu dois rentrer tôt; autrement, tu ne dîneras pas.*  
 “You must come back soon; otherwise, you’ll have no dinner.”

Second, French *autre*, like the adjective *différent*, and unlike Spanish *otro*, selects a so-called ‘item of comparison’ introduced by *que*:<sup>4</sup>

- (6) a. \**María se ha comprado otro vestido que Rosa.*  
 b. *Marie s’est achetée une autre robe que Rose.*  
 Lit.: “Mary has bought another dress than Rose.”
- (7) a. \**Juan ha leído otra novela que la que le recomendaron.*  
 b. *Jean a lu un autre roman que celui qui lui a été recommandé.*  
 Lit.: “John has read another novel than the one he was recommended.”

Third, if followed by an item of comparison, *autre* can function as a predicate in an attributive sentence:<sup>5</sup>

- (8) a. \**Mi hermano es otro de lo que parece.*  
 b. *Mon frère est autre que ce qu’il paraît.*  
 Lit.: “My brother is other than it seems.”
- (9) a. \**La situación es hoy otra de la ayer.*  
 b. *Aujourd’hui la situation est autre que celle d’hier.*  
 Lit.: “Today, the situation is other than yesterday’s.”

Finally, as shown in (10) and (11), French *autre*, when complemented by an item of comparison, may appear in postnominal position:<sup>6</sup>

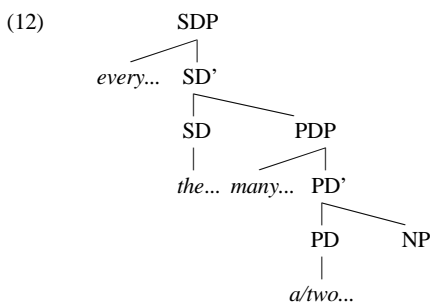
<sup>4</sup> Sentences can be found in Spanish in which *otro* seems to be selecting an item of comparison, for example, *No tengo otro sentimiento que éste* “I have no other feeling but this.” Notice, however, that negation is obligatory in these cases (cf. \**Tengo otro sentimiento que éste*), which shows that the *que* phrase is here a corrective phrase licensed by negation, and not the item of comparison of an adjective. In our view, the sequence *No...otro N que* “No...other N but” is then a discontinuous expression conveying an exclusiveness meaning (i.e., “I only have this feeling”).

<sup>5</sup> Sentences such as *La situación es otra* (Lit.: “The situation is other”) or *Mi problema es otro* (Lit.: “My problem is other”) are fine in Spanish, and might be taken as evidence for the adjectival status of Spanish *otro*. However, in Eguren and Sánchez (2003) it is shown that these are identifying, nonattributive clauses in which *otro* has a referential interpretation.

<sup>6</sup> For some of our French informants, *autre* can be both an attribute and a postnominal modifier

- (10) a. \**Ésa es una posición ideológica otra de la tuya.*  
 b. *Il s'agit d'une position idéologique autre que la tienne.*  
 Lit.: "That is an ideological stance other than yours."  
 (11) a. \**Nos hemos comprado un coche otro del tuyo.*  
 b. *Nous nous sommes achetés une voiture autre que la tienne.*  
 Lit.: "We have bought a car other than yours."

In order to account for this pattern of cross-linguistic variation, we will adopt the split DP-structure in (12) that is argued for in Zamparelli (2000), who follows a well-founded tradition (see recently Vangsnes 2001 and the references therein):



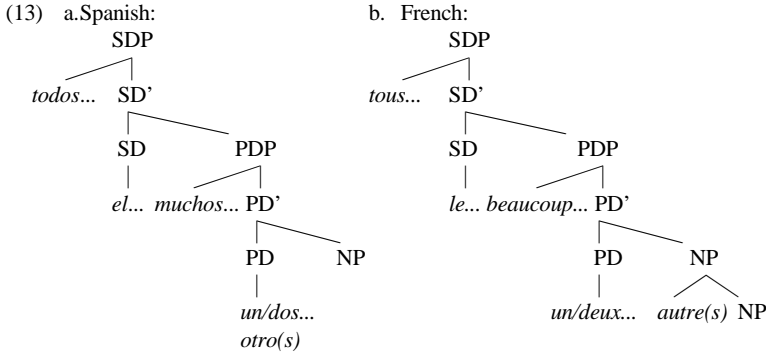
In (12), strong and weak determiners are generated at two different levels within the ‘determiner system’: a Strong Determiner Phrase (SDP) that denotes an individual, and a Predicative Determiner Phrase (PDP) that corresponds to Number Phrase or Quantifier Phrase in previous proposals, and denotes a property that is predicated of the head SD. A DP that includes the topmost level is an argument nominal. In predicate nominals, the topmost level is missing. Zamparelli (2000) also argues that indefinites with a strong reading move to the SDP, whereas indefinites with a weak reading remain in situ. As depicted in (12), Zamparelli finally holds that the heads of both the SDP and the PDP are filled by determiners that may have a ‘referential’ reading—in the sense of Fodor and Sag (1982)<sup>7</sup>—whereas the specifier of both projections is occupied by determiners that always have quantificational force (the idea that cardinal and vague numerals in particular must be generated in different positions is also pointed out in Verkuyl 1981 and Giusti 1992, among others).

---

without an item of comparison, particularly so if it is preceded by the intensifier *tout*, and focused, as in *Il s'agit d'une position idéologique TOUT AUTRE*.

<sup>7</sup> A DP headed by a determiner with a ‘referential’ reading is interpreted as a proper noun or a personal pronoun, and so lacks quantificational force. DPs with a referential reading can escape from a scope island (see (14b)), and do not show weak crossover effects.

Keeping these ideas in mind, we claim that Spanish *otro* is the head of a PDP, and so competes with the indefinite article and cardinal numerals for the very same position in the structure, whereas French *autre* is generated in a lower position, arguably as an NP adjunct. This is represented in (13):



By assigning these different structural positions to *otro* and *autre*, we can first give an account for the contrast we have seen so far: The predicate of contrast only has adjectival properties in French, whereas *otro* is both a prenominal predicate and a determiner in Spanish. And we also capture the fact that, like the indefinite article, Spanish *otro* is a weak determiner that may have a referential reading. As shown in (14), this is certainly the case: A DP headed by *otro* is allowed in presentational contexts, and can ‘escape’ from scope islands.

- (14) a. *Hay {un/otro} coche junto a la valla.*  
 “There is {a/another} car at the fence.”  
 b. *Si Fido ve a [{un/otro} chico que lo maltrataba]<sub>i</sub>, lo, muerde.*  
 “If Fido sees [{a/another} boy that tormented him], he bites him.”

Locating *otro* as the head of a PDP and *autre* as an NP adjunct finally allows us to explain their intricate combinatorial properties within the DP. As shown in (15)-(18), *otro* and *autre* behave alike in that they precede all kinds of prenominal adjectives, and follow both strong determiners and most weak determiners:<sup>8</sup>

<sup>8</sup> Spanish *otro* goes after vague numerals, but it also precedes them: *(Los) {muchos otros/otros muchos} libros* “(The) {many other/other many} books.” The second ordering follows from the fact that the indefinite determiners *muchos* “many” and *pocos* “few” can also function as prenominal adjectives (cf. *Los muchos e interesantes libros que he leído* Lit.: “The many and interesting books I have read”). For an in-depth analysis of these sequences see Eguren and Sánchez (2003).

- (15) a. *Los {otros prestigiosos/\*prestigiosos otros} profesores*  
 b. *Les {autres prestigieux/\*prestigieux autres} professeurs*  
 Lit.: “The {other prestigious/prestigious other} professors”
- (16) a. *{La/Esta} otra película*  
 b. *{L/Cet} autre film*  
 “{The/This} other film”
- (17) a. *{Algún/Ningún/Cualquier} otro libro*  
 b. *{Quelque/Aucun/N’importe quel} autre livre*  
 “{Some/No/Any} other book”
- (18) a. *{Varios/Bastantes/?Demasiados} otros libros*  
 b. *{Plusieurs/Assez d’/?Trop d’} autres livres*  
 “{Several/Enough/Too many} other books”

All these linear arrangements are straightforwardly derived from the structures in (13), with the natural additional assumption that adjectives with a ‘reference-related’ meaning are located in the highest position within the prenominal adjectival system.<sup>9</sup>

However, Spanish *otro* and French *autre* combine differently with both the indefinite article and cardinal numerals. On the one hand, Spanish *otro* and the indefinite article never co-occur, contrary to French, as shown in (19) and (20):

- (19) a. \**Un otro médico ha examinado al niño.*  
 b. *Un autre médecin a examiné l’enfant.*  
 “Another doctor has examined the child.”
- (20) a. \**Unos otros médicos han examinado al niño.*  
 b. *D’autres médecins ont examiné l’enfant.*  
 “Other doctors have examined the child.”

On the other hand, as illustrated in (21) and (22), *otro* goes before cardinal numerals, regardless of whether it is preceded by a strong determiner. French *autre* always follows cardinal numerals:

- (21) a. *{Otras dos/\*Dos otras} personas opinan lo mismo que yo.*  
 b. *{\*Autres deux/Deux autres} personnes sont de mon avis.*  
 “{Other two/Two other} people agree with me.”
- (22) a. *He traducido los {otros dos/\*dos otros} libros.*  
 b. *J’ai traduit les {\*autres deux/deux autres} livres.*  
 “I have translated the {other two/two other} books.”

---

<sup>9</sup> ‘Reference-related’ adjectives constitute a class of items that, by means of their inherent semantics, contribute to the identification of individuals. Spanish prenominal ‘determinative’ adjectives such as *mismo* ‘same,’ *distintos* ‘different,’ *propio* ‘own,’ and *demás* ‘the rest of’ belong to this class.

And much the same is true for cardinal anaphors. As shown in (23a), the Spanish word *tantos* may denote an exact quantity that anaphorically refers to a cardinal numeral and, in that case, it must follow indefinite *otros*. As can be seen in (23b), *autant* can also be a cardinal anaphor in French, but it precedes *autres*, as cardinal numerals do in that language:

- (23) a. *Yo he leído tres libros y Ana {otros tantos/\*tantos otros}.*  
 b. *J'ai lu trois livres et Anne {\*d'autres autant/autant d'autres}.*  
 "I have read three books and Ann has read another three."

In our view, the paradigms in (19) to (23) are related to the fact that *otro* is a weak determiner in Spanish, while *autre* is an adjective in French, and they directly result from linearizing the structures in (12) and (13), if we also take into account examples like those in (24) and (25), which show that in Spanish, but not in French, cardinal numerals can be generated in two different positions, either as the head of a PDP or in a lower position, closer to the noun:

- (24) a. *He traducido los dos {magníficos/primeros} libros de Cela.*  
 b. *J'ai traduit les deux {magnifiques/premiers} livres de Cela.*  
 "I have translated the two {excellent/first} books by Cela."  
 (25) a. *He traducido los {magníficos/primeros} dos libros de Rulfo.*  
 b. *?? J'ai traduit les {magnifiques/premiers} deux livres de Rulfo.*  
 "I have translated the {excellent/first} two books by Rulfo."

Our explanation for all these data goes as follows. Spanish *otro* cannot combine with the indefinite article just because both lexical items compete for the same position in the structure in (12), and the order <cardinal numeral + *otros*> is also ruled out for the very same reason. But the reversed order <*otros* + cardinal numeral> is well formed, and that is the case because cardinal numerals can be generated in Spanish in a position closer to the noun. *Autre*, instead, is a prenominal adjective in French, and may then follow both the indefinite article and cardinal numerals. But it cannot precede cardinal numerals, because these are never generated in a position closer to the noun in French.

Let us now sketch the picture in mixed systems, such as those of Italian, Catalan, and Portuguese, in which the predicate of contrast can be both a determiner and a prenominal adjective. The basic paradigms in (26)-(31) correspond to some of the French and Spanish data we have already gone through. As shown in (26)-(28), in Italian: (a) singular *altro* can never be a determiner, but plural *altri* does license a nominal expression in an argument position (cf. (26));<sup>10</sup> (b) as we predict, the predicate of contrast retains some

<sup>10</sup> This is a common paradigm within the adjectival determiner system across languages, which

adjectival properties, but loses others (e.g., it is the input for the formation of a *-mente* adverb, but it does not select an item of comparison; cf. (27)); and (c) it combines with the indefinite article, and may be either preceded or followed by cardinal numerals (cf. (28)).

- (26) a. \**Altro studente ha superato l'esame.*  
 "Another student has passed his exam."  
 b. *Altri studenti hanno superato l'esame.*  
 "Other students have passed their exam."
- (27) a. *Lo voglio fare altrimenti.*  
 "I want to do it in another way."  
 b. \**Maria si è comprata un altro vestito che Rosa.*  
 Lit.: "Mary has bought another dress than Rosa."
- (28) a. *Un altro medico ha esaminato il bambino.*  
 "Another doctor has examined the child."  
 b. {*Altre due/Due altre*} *persone sono della mia opinione.*  
 "{Other two/Two other} people agree with me."

Catalan is very much like Italian in this respect. As can be seen in (26')-(28'), which correspond to the sentences in (26)-(28), Catalan *altre* can be a determiner only in the plural form (cf. (26')), has some adjectival properties (e.g., unlike Italian *altro*, the adverb *altrament* is barely used,<sup>11</sup> but Catalan *altre* does select an item of comparison; cf. (27')), combines with the indefinite article, and can both precede and follow cardinal numerals (cf. (28')):

- (26') a. \**Altres estudiant ha passat l'examen.*  
 b. *Altres estudiants han passat l'examen.*
- (27') a. \**Jo ho vull fer altrament.*  
 b. *La Maria s'ha comprat un altre vestit que la Rosa.*
- (28') a. *Un altre metge ha vist el nen.*  
 b. {*Altres dues/Dues altres*} *persones opinen el mateix que jo.*

In Portuguese, as illustrated in (29)-(31): (a) *outro*, like Spanish *otro*, can be a determiner in both the singular and plural forms (cf. (29)); (b) *outro* does not have the adjectival properties of French *autre* (cf. (30)); but (c), unlike Spanish *otro*, *outro* can combine with the indefinite article (conveying a contrastive

---

shows that the plural formal feature allows an adjective with a referential/quantificational-like meaning to become a determiner. The Spanish adjective *distinto* "different," for example, like Italian *altro*, licenses a nominal expression in an argument position only in the plural form: For example, *Distintas personas piensan como yo* "Different people agree with me" versus \**Distinta persona piensa como yo* "\*Different person agrees with me."

<sup>11</sup> The Catalan adverb *altrament* is only used as a discourse marker, as in *Has de tornar aviat; altrament non soparàs* "You must come back soon; otherwise, you will not have dinner."





b. *J'ai lu un autre livre.*

"I have read {a different book/one more book}."

This seems to be a natural semantic extension once we realize that an implicit quantification over individuals is being carried out when a distinct member of an identical class is identified. However, it is the case that only French *autre* (with an additive meaning) turns out to be an actual degree operator that heads an additive construction with a referential base.

Additive constructions are headed by a dyadic operator that selects two magnitudes: Let us call them the 'base' and the 'differential.' Following Brucart (2003), two types of additive constructions are to be distinguished in Spanish: additives with a quantitative base introduced by the preposition *de*, as in (34a), and additives with a referential base beginning with the conjunction *que*, as in (34b).

(34) a. *Tengo algo más de cuatro euros en el bolsillo.*

Lit.: "I have something more of four euros in my pocket."

b. *Tengo algo más que cuatro euros en el bolsillo.*

Lit.: "I have something more than four euros in my pocket."

In both cases, the degree operator *más* adds the denotation of the differential (the indefinite *algo* "something") to the denotation of the base (*cuatro euros* "four euros"). But in the additive construction with a quantitative base in (34a), two quantities are added up, and the cardinality of a set is obtained. And, consequently, we understand that I have an amount of euros higher than four in my pocket. However, in the additive construction with a referential base in (34b), it is now two sets that are added up, resulting in a larger set that puts together the members of both sets. That is why we understand that, in this case, I have four euros and something else in my pocket.

As shown in (35), additives with a quantitative base are headed by the additive operators *más* and *plus* in Spanish and French, and both *otro* and *autre* are ruled out:

(35) a. *Compré dos libros {más/\*otros} de los cuatro previstos.*

b. *J'ai acheté deux livres en plus des quatre que j'avais prévus.*

c. *\*J'ai acheté deux livres d'autres des quatre que j'avais prévus.*

Lit.: "I got two books {more/other} of the four I was to buy."

However, as illustrated in (36) and (37), additives with a referential base can be headed by *autre* in French, but not by *otro* in Spanish:<sup>13</sup>

<sup>13</sup> In all the French examples in (36) to (39) a contrastive reading is obtained as well, and the

- (36) a. *Conozco {más/\*otros} filósofos que Descartes.*  
 b. *Je connais d'autres philosophes que Descartes.*  
 Lit.: "I know of {more/other} philosophers than Descartes."  
 (37) a. *He leído {más/\*otros} libros que "Rayuela."*  
 b. *J'ai lu d'autres livres que "Rayuela."*  
 Lit.: "I have read {more/other} books than 'Rayuela.'"

There is then a second locus for variation, which also shows up if the sentences in (36) and (37) are negated, giving rise to so-called 'restrictive constructions,' in which a set is strictly limited instead of being expanded, as in (38)-(39):<sup>14</sup>

- (38) a. *No conozco {más/\*otros} filósofos que Descartes.*  
 b. *Je ne connais pas d'autres philosophes que Descartes.*  
 Lit.: "I don't know of {more/other} philosophers than Descartes."  
 (39) a. *No he leído {más/\*otros} libros que "Rayuela."*  
 b. *Je n'ai pas lu d'autres livres que "Rayuela."*  
 Lit.: "I have not read {more/other} books than 'Rayuela.'"

As can be seen in (40)-(45), the differential in additive and restrictive constructions may also appear to the left of French *autre*. In that case, the differential can be an indefinite pronoun, like *quelque chose* in (40b) or *quelqu'un* in (41b); an interrogative pronoun, like *quoi* in (42b) or *qui* in (43b); or a negative indefinite, like *rien* in (44b) or *personne* in (45b):

- (40) a. *Compré algo {más/\*otro} que lo que pensaba comprar.*  
 b. *J'ai acheté quelque chose d'autre que ce que je pensais acheter.*  
 Lit.: "I bought something {more/other} than what I was to buy."  
 (41) a. *Lo sabe alguien {más/\*otro} que tú.*  
 b. *Quelqu'un d'autre que toi le sait.*  
 Lit.: "Somebody {more/other} than you knows it."  
 (42) a. *¿Qué {más/\*otro} compraste que lo que pensabas comprar?*  
 b. *Quoi d'autre as-tu acheté que ce que tu pensais acheter?*  
 Lit.: "What {more/other} did you buy than what you were to buy?"  
 (43) a. *¿Quién {más/\*otro} que tú lo sabe?*  
 b. *Qui d'autre que toi le sait?*  
 Lit.: "Who {more/other} than you knows it?"

---

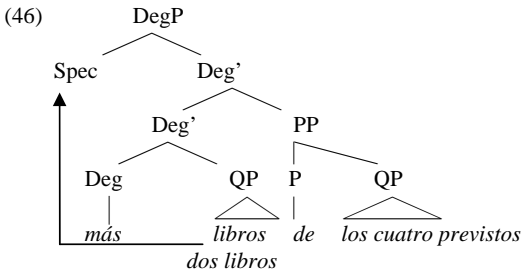
operator *plus de* "more" can be used instead of additive *autres*. Notice also that Spanish *otro(s)* with an additive interpretation can co-occur with a phrase introduced by the expression *además de* "apart from": *He leído otros libros, además de "Rayuela"* "I have read other books apart from 'Rayuela.'" In our view, this phrase is an adjunct, and not a selected complement, as shown by its free linear arrangement: *Además de "Rayuela," he leído otros libros; He leído, además de "Rayuela," otros libros.*

<sup>14</sup> The Spanish examples in (38a) and (39a) are well formed with singular *otro*, but see fn. 4.

- (44) a. *No compré nada {más/\*otro} que lo que pensaba comprar.*  
 b. *Je n'ai acheté rien d'autre que ce que je pensais acheter.*  
 Lit.: "I did not buy anything {more/other} than what I was to buy."  
 (45) a. *No lo sabe nadie {más/\*otro} que tú.*  
 b. *Personne d'autre que toi ne le sait.*  
 Lit.: "Nobody {more/other} than you knows it."

In the Spanish sentences in (40)-(45), the operator *más* is obligatorily used, and *otro* is ruled out. To summarize, then, both Spanish *otro* and French *autre* may have an additive interpretation that results from their basic contrastive meaning, but only French *autre* has the properties of a degree operator that projects an additive construction with a referential base.

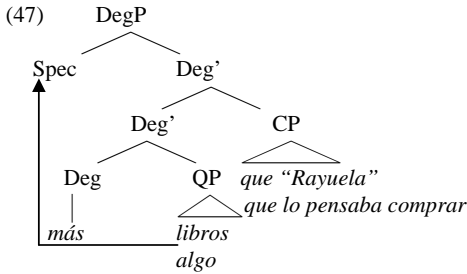
In order to capture this pattern of microvariation configurationally, we assume the ideas in Brucart (2003), who extends to additives with a quantitative base a previous proposal by Sáez del Álamo (1997) for standard comparatives, arguing that additives with a quantitative base are to be analyzed as in (46):



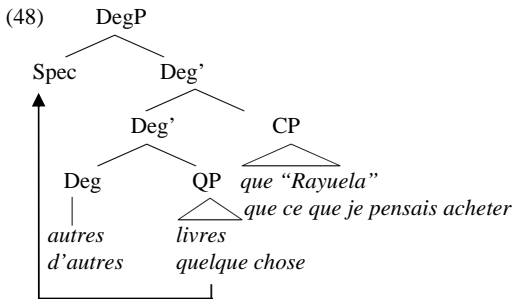
These constructions are then DegPs (see, e.g., Abney 1987; Corver 1991, 1997), which are headed by an additive operator that selects two complements: a differential that takes the form of a phrase denoting some quantity, and a base corresponding to a PP that contains another QP. According to Brucart, in (46) a suitable item has to move to the specifier of DegP in order to check a [+differential] feature belonging to the degree word that heads the phrase. The whole QP will move if the differential has a non-null head, as in the phrase *dos libros más de los cuatro que pensaba comprar* (Lit.: "two books more of the four I was meant to buy"), and whenever the differential is a bare plural, an empty operator will do the job at LF, as in the sequence *más libros de los cuatro que pensaba comprar* (Lit.: "more books of the four I was meant to buy").

As depicted in (47), this analysis can be easily extended to additive

constructions with a referential base headed by the operator *más* in Spanish, with two provisos: The differential now has to be a QP with an existential interpretation, and the base takes the form of a CP containing a set of individuals that is a subset of the class denoted by the differential. As in (46), in (47) either an empty operator or the whole QP moves to the specifier of DegP, and so checks the [+differential] feature of the Deg head.



Moreover, as shown in (48), the same proposal applies to French additives with a referential base, the only difference being that *autre* is an additive operator that selects both a differential and a base, and so heads a DegP:



To conclude, let us briefly examine the cases of Italian, Catalan, and Portuguese with respect to this second locus for variation. Both Portuguese and Catalan are now just like Spanish.<sup>15</sup> As illustrated in (49) and (49'), *outro* and *altre* can select neither a base nor a differential, and the degree operator *mais/més* obligatorily shows up in all additive constructions:

(49) a. *Comprei algo {mais/\*outro} do que o que pensava comprar.*

<sup>15</sup> Romanian *alt* is not an additive operator either.

- Lit.: “I bought something {more/other} than what I was to buy.”
- b. *Quem {mais/\*outro} que tu o sabe?*  
 Lit.: Who {more/other} than you knows it?”
- (49') a. *Vaig comprar alguna cosa {més/\*altra} del que pensava comprar.*  
 b. *Qui {més/\*altre} ho sap, apart de tu?*

In Italian, on the other hand, we come across a kind of mixed system again. As shown in (50), *altro*, with an additive interpretation, is a defective degree operator, as it can be preceded by an indefinite or an interrogative pronoun, but it does not select an additive base:

- (50) a. *Ho comprato qualcos'altro (\*che ciò che pensavo di comprare).*  
 b. *Chi altro (\*che te) lo sa?*

### 3. *Otro in Old Spanish and dialectal variation*

In this last section we will offer some Old Spanish and American Spanish data that give support to the idea that microparameters are a privileged target for historical change and dialectal variation. Let us begin with Old Spanish.

To a limited extent, Old Spanish *otro* behaved as an additive operator, and it clearly had the adjectival properties of French *autre*, but it was also beginning to develop determiner-like properties. We will go through these two patterns in turn.

As illustrated in (51),<sup>16</sup> and in clear contrast with the situation today, some data can be found in Medieval and Classical Spanish in which *otro* seems to be heading an additive construction. In that case, the differential could be a negative or an interrogative indefinite (cf. (51a,b)), and examples are also attested in which both the differential and the base are overt (cf. (51c)):

- (51) a. *Puede ella no perderse y no puede nadie otro perdella.*  
 Lit.: “She may not be ruined, and nobody other can ruin her.” (CORDE, 1598)
- b. *No están obligados a ayunar los que no tienen veinte y un años.*  
 P: *¿Y quiénes otros?* R: *Los que no pueden ayunar cómodamente.*  
 Lit.: “It is not obligatory for those under twenty-one to fast.  
 Q: And for who others? A: For those who cannot fast comfortably.”  
 (CORDE, 1591)
- c. *De hecho, ¿quién otro que Hardyl, que casi toda su vida...?*  
 Lit.: “In fact, who other than Hardyl, who almost all his life...?” (CORDE, 1786)

These data suggest that *otro* may have had the properties of a degree operator

<sup>16</sup> The Old Spanish and American Spanish examples in (51)-(60) are taken from the CORDE (Diachronic Reference Corpus of Spanish) and the CREA (Reference Corpus of Current Spanish).

in previous stages of the language, though perhaps not in any systematic way, as it does in French. As we saw in section 2, such a microparametric option is not active in modern Standard Spanish.

The adjectival status of *otro* in Medieval and Classical Spanish is much more conspicuous. First, *otro* very often followed the indefinite *un*, as in (52), but it could also license a nominal expression in an argument position (53). Second, as illustrated in (54), *otro* followed cardinal numerals. However, the reversed order <*otro* + cardinal numeral> is also attested, and was more frequent.

- (52) a. *Un ombre que auia uinas se aueno con un otro que auia oueias...*  
 "A man who had vineyards agreed with another who owned sheep."  
 (CORDE, 1250-1300)
- b. *...uino a unos otros pueblos asaç habundantes de lauores.*  
 "...he came to some other villages with plenty of work." (CORDE, 1385)
- (53) a. *... que otro su pariente no pueda desafiar por ellos.*  
 Lit.: "that other their relative could not defy in their place." (CORDE, 1348)
- b. *... pero que avn el se tenje otros pocos de djneros para despende.*  
 "... but that he still had other few coins to waste." (CORDE, 1300-1305)
- (54) a. *... e despues le dio dos otros golpes...*  
 "... and then he gave him two other blows..." (CORDE, 1376-1384)
- b. *... dio de nuevo otros tres por amor de las virgenes.*  
 "... (s)he gave again other three for love to the virgins." (CORDE, 1594)

Finally, Old Spanish *otro* had some clear-cut French-like adjectival properties (see section 1), that are lost nowadays:

(i) The predicate of contrast was the input for the formation of the adverb *otramente* "otherwise" (notice that, most significantly, speakers stopped using both this adverb and the sequence *un otro* "another" at approximately the same time), as in (55):<sup>17</sup>

- (55) *Esto es neçessario en Dios ca otramente non seria infinitamente perfecto.*  
 "This is necessary in God for otherwise he will not be infinitely perfect." (CORDE, 1437)

(ii) It selected an item of comparison, as in (56):

- (56) *... el enperador hi enuio mas gent, otra que la primera de que era capitan.*  
 Lit.: "The emperor sent more people there, other than the former he was the

---

<sup>17</sup> In Old Spanish, compounds including adjectival *otro* such as *otrosí* "in other ways" or *otrora* "in other times" were also used. These forms were lost. However, not surprisingly (see section 1), the Italian compounds *altrove* "in another place" and *altronde* "from another place" are still used.

captain of.” (CORDE, 1377)

(iii) If followed by an item of comparison, it could also function as a predicate in an attributive sentence, as in (57):

(57) ... y assi les quedo tal nombre, con la pronunciación castellana que es otra que la italiana.

Lit.: “... and so they got that name, with the Castilian pronunciation which is other than the Italian one.” (CORDE, 1529)

To summarize so far, what we find in Old Spanish is some kind of mixed system, like those of Italian, Catalan, and Portuguese (see section 1): In Medieval and Classical Spanish *otro* was mainly an adjective, but it was also developing determiner-like properties. This process has come to its end in modern Standard Spanish. Nowadays, as we have shown, apart from being a pronominal predicate of contrast, *otro* lacks all its former adjectival properties, and clearly belongs to the determiner word class.

Let us now finally have a look at dialectal variation within Spanish on the topic at hand. In this respect, there are relevant data that seem to indicate that at least some varieties of American Spanish retain the operator and adjectival-like properties *otro* had in Old Spanish.

As shown in (58), examples can be found in American Spanish in which *otro* (with an additive interpretation) is preceded by an indefinite or an interrogative pronoun:

(58) a. *Luego me echarás de menos y besos como los míos, nadie otro.*

Lit.: “Then you will miss me and kisses like mine, nobody other.”  
(CREA, Colombia)

b. *Bueno, y ¿quién otro? Luis Spinetti Dini...estaba también.*

Lit.: “Well, and who other? Luis Spinetti Dini...was there also.”  
(CREA, Venezuela)

But, once again, the adjectival, French-like properties *otro* has in current American Spanish are much better attested. In (at least some varieties of) American Spanish, *otro* can co-occur with the indefinite *un* (cf. (59a-b)), and may follow cardinal numerals (cf. (59c-d)):

(59) a. *Días antes de su muerte, un otro golpe hirió su corazón.*

“Some days before he died, another blow touched his heart.” (CREA, Bolivia)

b. *Ambos equipos se enfrentaron una otra vez desde esa gran final.*

“Both teams met once again since that great final.” (CREA, Costa Rica)

c. *Se comprometieron a comprar un avión 737 y a alquilar dos otros.*

- “They agreed to buy a 737 and to hire another two.” (CREA, Bolivia)  
 d. *Los dos otros tubos...tienen respectivamente...*  
 “The two other pipes...have respectively...” (CREA, Venezuela)

Furthermore, (a) the adverb *otramente* “otherwise” is still used (cf. (60a)), (b) the predicate of contrast selects an item of comparison (cf. (60b)), and (c), in that case, it can be the predicate in an attributive clause (cf. (60c)):

- (60) a. *Entre tu soledad y la mía hay una gran diferencia, otramente importante: a ti te sirve.*  
 “Between your solitude and mine there is a big difference, otherwise important: It is useful to you.” (CREA, Ecuador)  
 b. *...habría una otra Bolivia económica que la que conocemos hoy.*  
 Lit.: “...There would be another economic Bolivia than today’s.” (CREA, Bolivia)  
 c. *...esta gente, contra la opinión general, que es otra que la mía, es de un agradecimiento eterno.*  
 Lit.: “... these people, against the common view, which is other than mine, are most grateful.” (CREA, Perú)

## REFERENCES

- Abney, Stephen. 1987. *The English Noun Phrase in its Sentential Aspect*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Baker, Mark C. 1996. *The Polysynthesis Parameter*. Oxford: Oxford University Press.
- Brucart, José María. 2003. “Adición, sustracción y comparación: un análisis composicional de las construcciones aditivo-sustractivas del español”. *Actas del XXIII CILFR* ed. by Fernando Sánchez Miret, vol. I, 11-60. Tübingen: Niemeyer.
- Corver, Norbert. 1991. “Evidence for DegP”. *NELS* 21.33-47.
- , 1997. “Much-Support as a Last Resort”. *Linguistic Inquiry* 28.119-164.
- Diachronic Reference Corpus of Spanish (CORDE). <http://www.rae.es/cordenet.html> (January 2003)
- Eguren, Luis & Cristina Sánchez. 2003. “La gramática de *otro*”. *Revista Española de Lingüística* 33.66-122.
- Fodor, Janet Dean & Ivan A. Sag. 1982. “Referential and Quantificational Indefinites”. *Linguistics and Philosophy* 5.355-398.
- Giusti, Giuliana. 1992. *La Sintassi dei Sintagma Nominali Quantificati: uno studio comparativo*. Ph. D. dissertation, University of Venice.



- Holmberg, Anders & Görel Sandström. 1996. "Scandinavian Possessive Constructions from a Northern Swedish Viewpoint". *Microparametric Syntax and Dialect Variation* ed. by James R. Black & Virginia Motapanyane, 95-120. Amsterdam & Philadelphia: John Benjamins.
- Kayne, Richard. 1996. "Microparametric Syntax. Some Introductory Remarks". *Microparametric Syntax and Dialect Variation* ed. by James R. Black & Virginia Motapanyane, ix-xviii. Amsterdam & Philadelphia: John Benjamins.
- Reference Corpus of Current Spanish (CREA). <http://www.rae.es/creanet.html> (January 2003)
- Sáez del Álamo, Luis A. 1997. "Cuantificadores y sintagmas de medida". *Proceedings of the IV Colloquium on Generative Grammar* ed. by Natàlia Català & Maria Bargalló, 162-186. Tarragona: Universitat Rovira i Virgili.
- Vangsnes, Ø. Alexander. 2001. "On Noun Phrase Architecture, Referentiality, and Article Systems". *Studia Lingüística* 55.249-299.
- Van Peteghem, Marleen. 1997a. "Autre et ses correspondants en roumain". *Revue Romane* 32.27-50.
- , 1997b. "Mécanismes anaphoriques sous-jacents aux 'indéfinis' autre et même". *Relations anaphoriques et (in)cohérence* ed. by Walter De Mulder, Liliane Tasmowski-DeRyck, & Carl Veters, 187-200. Amsterdam: Rodopi.
- Verkuyl, Henk J. 1981. "Numerals and Quantifiers in X-bar Syntax and Their Semantic Interpretation". *Formal Methods in the Study of Language* ed. by Jeroen A. G. Groenendijk, Theo M. V. Janssen, & Martin J. B. Stokhof, 567-599. Amsterdam: Mathematisch Centrum.
- Zamparelli, Roberto. 2000. *Layers in the Determiner Phrase*. New York: Garland.

## ON THE STRUCTURE OF SYNCRETISM IN ROMANIAN CONJUGATION

RONALD F. FELDSTEIN  
*Indiana University*

### 0. *Introduction*

Romanian conjugation displays several cases of syncretism, in which two paradigmatic slots share the same grammatical desinence.<sup>1</sup> This syncretism occurs in both the present and imperfect tenses of the indicative, as well as in the subjunctive. In the present tense, syncretism occurs in the 1sg/3pl, the 3sg/3pl, and the 1sg/2sg forms; in the imperfect, the 1sg/1pl forms are always syncretic; number syncretism is always found in the 3sg/3pl of the subjunctive. I will claim that all such instances of syncretism should be divided into two categories, based on whether they are phonologically conditioned or not. When syncretism is phonologically conditioned, it will only apply to a specific phonological subset of verbs, while the nonphonological type applies across the board. In Romanian, the present-tense varieties of syncretism can be explained by phonological conditioning, once the appropriate underlying morphophonemic constructs and rules are established. On the other hand, the syncretisms of the imperfect and subjunctive are not phonologically conditioned and, as such, apply to every verb without exception.

Two recent publications attempt to analyze the syncretisms of Romanian conjugation: sections of a book by Stump (2001:213-215) and a paper by Bobaljik (2002:65-66). This chapter will show that both authors operate with structurally inadequate models of Romanian conjugation and syncretism, which leads them to make analytical errors. Their errors are mainly due to the fact that they do not distinguish phonologically conditioned, or variable syncretisms, from invariant syncretisms that have nothing to do with a

---

<sup>1</sup> This chapter represents Romanian sounds by means of orthographic symbols. Letters that should be specially noted are *ă* (equivalent to the mid central vowel *schwa*, i.e., [ə]), and *î*, which represents the high central vowel [i]. The high vowels *i* and *u*, in postvocalic position (e.g., *dai*, *dau*) represent nonsyllabic [i] and [u]. The sequence *ea* represents the diphthong [e̯a]. The consonants *ș*, *ț* represent [ʃ] and the affricate [ts], respectively.

particular phonological environment. These inaccuracies stem from the authors' reliance on textbook notions of the various subtypes of Romanian syncretism, which are traditionally, but incorrectly, said to be determined by conjugation type. Since these are not really errors made by Stump and Bobaljik *per se*, but are caused by the inadequacies of the traditional treatment of Romanian conjugation and syncretism, it becomes clear that the traditional approach needs to be modified.

This chapter will begin with a brief review of the basic facts about Romanian syncretism (section 1), where the traditional conjugations will be compared to the actual distribution of syncretic types. Next (section 2), a new segmentation and morphophonemic analysis of Romanian conjugation is offered, which resolves some of the difficult issues of Romanian syncretism. Phonologically conditioned syncretism will be treated as underlyingly non-syncretic; that is, it should be considered to be a superficial, or surface, syncretism, while the nonphonological variety remains as a deeper syncretism. Section 3 examines how Stump (2001) and Bobaljik (2002) rely on the textbook view of conjugation for their analyses of Romanian syncretism, and the flaws of this approach are pointed out. While some of these errors are due to not considering all the data, others stem from an inefficient segmentation of the verb stem and desinence. Section 4 looks at the issue of whether a surface perspective can offer any useful linguistic generalizations about Romanian syncretism.

Since the work of Stump (2001) and Bobaljik (2002) is relevant to this chapter only insofar as it illustrates the inadequacy of the traditional textbook approach to present-tense syncretism, as based on conjugational type, rather than stem phonology, I will not specifically comment on the other main issue they discuss—whether these syncretisms should be treated as rules of referral<sup>2</sup> (the notion that one paradigmatic cell is replaced with the other) or strictly as rules of impoverishment (the idea that both cells are neutralized with respect to a particular grammatical feature). Generally speaking, Stump opts for the use of rules of referral, while Bobaljik prefers to operate with the method of impoverishment. However, I would emphasize that one cannot even begin to make linguistic judgements about the merits of impoverishment versus referral without a more accurate notion of the structural units of Romanian syncretism.

---

<sup>2</sup> Bobaljik (2002:66) maintains that impoverishment can handle all of the cases of Romanian syncretism and that rules of referral (introduced by Zwicky 1985:372) are too powerful and unrestricted a mechanism. Stump (2001) opts for unstipulated syncretism (equivalent to impoverishment) in some instances, but posits rules of referral in others.

### 1. *Brief review of Romanian syncretism*

I maintain that present-tense syncretisms are a function of the phonological properties of the verb stem. The traditional system, as repeated by Stump (2001) and Bobaljik (2002), states that there are four basic conjugations, and that each conjugational type utilizes a different set of desinences (and syncretisms), unrelated to phonology. I claim that each of the traditional conjugations merely refers to a particular stem-final theme vowel, which can be considered a verbal formant. Furthermore, I propose that the underlying desinences are identical for all four of the so-called conjugations, and that the surface differences of conjugation are the result of the phonological interaction of the different stem-final themes with the unified set of desinences. In many cases, the traditional system of conjugation does a poor job of capturing what actually occurs, especially in the area of syncretism.

Table 1 lists the four traditional conjugational types. Type I has the theme vowel *-a* and usually has 3sg/3pl syncretism; types II, III, and IV have the theme vowels *-e*, *-ea*, *-i*, (and *-î*<sup>3</sup>) and most frequently have 1sg/3pl syncretism. Table 2 presents present-tense paradigms for these and other verbs.

Conj.	Stem type (infinitive)	Examples
I	<i>-a</i>	<i>invita</i> “invite,” <i>tăia</i> “cut,” <i>sufla</i> “breathe,” <i>da</i> “give”
II	<i>-ea</i>	<i>tăcea</i> “be silent,” <i>umplea</i> “fill (variant),” <i>bea</i> “drink”
III	<i>-e</i>	<i>bate</i> “beat,” <i>umple</i> “fill”
IV	<i>-i</i>	<i>sări</i> “jump,” <i>coborî</i> , “descend,” <i>sui</i> “climb,” <i>suferi</i> “suffer,” <i>şti</i> “know”

Table 1: *Traditional Romanian conjugations (Graur 1966:246)*

Infinitive	Conj. type	1sg	2sg	3sg	1pl	2pl	3pl
<i>invita</i>	I	<i>invít</i>	<i>invíţi</i>	<b><i>invită</i></b>	<i>invităm</i>	<i>invitaţi</i>	<b><i>invită</i></b>
<i>tăia</i>	I	<b><i>tái</i></b>	<b><i>tái</i></b>	<b><i>táie</i></b>	<i>tăiém</i>	<i>tăiaţi</i>	<b><i>táie</i></b>
<i>sufla</i>	I	<i>súflu</i>	<i>súfli</i>	<b><i>súflă</i></b>	<i>suflám</i>	<i>suflaţi</i>	<b><i>súflă</i></b>
<i>da</i>	I	<b><i>dáu</i></b>	<b><i>dái</i></b>	<b><i>dă</i></b>	<i>dám</i>	<i>dáţi</i>	<b><i>dáu</i></b>
<i>tăcea</i>	II	<b><i>tác</i></b>	<i>táci</i>	<i>táce</i>	<i>tácém</i>	<i>táceţi</i>	<b><i>tác</i></b>
<i>umplea</i>	II	<b><i>úmplu</i></b>	<i>úmpli</i>	<i>úmples</i>	<i>úmplém</i>	<i>úmpléţi</i>	<b><i>úmplu</i></b>
<i>bea</i>	II	<b><i>beáu</i></b>	<i>béi</i>	<i>beá</i>	<i>bém</i>	<i>béţi</i>	<b><i>beáu</i></b>
<i>báte</i>	III	<b><i>bát</i></b>	<i>báţi</i>	<i>báte</i>	<i>bátem</i>	<i>báteţi</i>	<b><i>bát</i></b>
<i>úmples</i>	III	<b><i>úmplu</i></b>	<i>úmpli</i>	<i>úmples</i>	<i>úmplem</i>	<i>úmplesţi</i>	<b><i>úmplu</i></b>
<i>sári</i>	IV	<b><i>sár</i></b>	<i>sári</i>	<i>sáre</i>	<i>sárim</i>	<i>sáriţi</i>	<b><i>sár</i></b>
<i>coborî</i>	IV	<i>cobór</i>	<i>cobóri</i>	<b><i>coboáră</i></b>	<i>coborîmi</i>	<i>coborîţi</i>	<b><i>coboáră</i></b>
<i>sui</i>	IV	<i>súi</i>	<i>súi</i>	<b><i>súie</i></b>	<i>súim</i>	<i>súiş</i>	<b><i>súie</i></b>
<i>suferi</i>	IV	<i>súfăr</i>	<i>súferi</i>	<b><i>súferă</i></b>	<i>suferîm</i>	<i>suferîţi</i>	<b><i>súferă</i></b>
<i>şti</i>	IV	<b><i>ştíu</i></b>	<i>ştíi</i>	<b><i>ştíe</i></b>	<i>ştím</i>	<i>ştíţi</i>	<b><i>ştíu</i></b>

Table 2: *Sample conjugations of verbs in Table 1, with syncretic cells marked in bold*

<sup>3</sup> Traditional grammar places theme vowels *-i* and *-î* in conjugation IV.

Table 2 shows that all of the verbs of a given conjugation type (i.e., I, II, III, IV) do not necessarily share the same pattern of syncretism. The following two reasons show why it is therefore wrong to maintain the traditional conjugation types when speaking of syncretism:

(i) Verbs of the same traditional conjugation type (i.e., with the same theme vowel) can have different patterns of syncretism. For example, *invita* “invite” has syncretism of the 3sg/3pl type. However, *tăia* “cut” has both the 1sg/2sg and 3sg/3pl types, while *da* “give” has the 1sg/3pl type, in spite of the fact that all of these stems end in the theme vowel *-a*. Although these verbs are ostensibly all of the same type, their different types of syncretism are due to the fact that the theme is not monosyllabic in the first two verbs, but is monosyllabic in the third (*da*); the monosyllabic stem causes the theme vowel to be stressed in all present forms, and conditions syncretism of the 1sg/3pl type, in spite of the *-a* theme. Furthermore, the second verb (*tăia*) has [j̥] (‘yod’) as its stem-final consonant, which causes an additional 1sg/2sg syncretism. In other words, phonology, rather than conjugational class, causes the specific pattern of present-tense syncretism.

(ii) Verbs of different traditional conjugations, often assumed to have different syncretic patterns, can actually have the same type of syncretism, if certain phonological conditions are met. For example, type I *invita* and type IV *sui* and *coborî* all share syncretism of the 3sg/3pl type. Phonologically, this is conditioned by the fact that all have back vowels either as the theme or immediately preceding the theme vowel, in spite of the fact that they do not belong to the same traditional conjugational classes.

In other words, the patterns of present-tense syncretism are predictable on the basis of the phonological shape of the stem, rather than the traditional conjugation class, so it is incorrect to identify particular syncretic types with the traditional conjugations. Table 3 demonstrates that each major type of present-tense syncretism can be correlated with a variety of different traditional conjugation types (i.e., theme-vowel classes).

For example, Table 3 shows that syncretism of the 1sg/3pl type can be manifested by verbs of all four conjugation types. Furthermore, each concrete manifestation of syncretism can be correlated to specific phonological properties of the stem. As illustrated, the 1sg/3pl syncretic form can end in the nonsyllabic glide [ɥ], vocalic [u], or a consonant. In the first instance, the stem is monosyllabic and can have a theme vowel that cuts across the broad range of traditional conjugations I, II, and IV. These cases are not isolated exceptions, but defined phonological types.

Syncretic cells	Phonological manifestation of syncretisms
1sg/3pl present	<p>1. Syncretic form ends in non-syllabic glide <math>-u</math> (theme vowels <math>-a, -ea, -i</math>): Monosyllabic stems, e.g., 1sg/3pl <i>dau, beau, știu</i> (infin. <i>da/bea,ști</i>)</p> <p>2. Syncretic form ends in the vowel <math>-u</math> (theme vowels <math>-ea, -e</math>): Front-vowel theme preceded by [obstruent + liquid] cluster, e.g., 1sg/3pl <i>umplu</i> (infin. <i>umplea ~ umple</i>)</p> <p>3. Syncretic form ends in zero (followed by pretheme consonant; theme vowels <math>-ea, -e, -i</math>): Non-monosyllabic stems, e.g., 1sg/3pl <i>tac, bat, sar</i> (infin. <i>tăcea/bate/sări</i>)</p>
3sg/3pl present	<p>1. Syncretic form ends in <math>-ă</math> (theme vowels <math>-a, -i, -î</math>): a. Non-monosyllabic stem in non-front vowel <math>-a</math> or <math>-î</math>, preceded by nonpalatal, e.g., 3sg/3pl <i>invită, coboară</i> (infin. <i>invita, coborî</i>) b. Polysyllabic stems in theme vowel <math>-i</math>, where present-tense stress falls two syllables before the theme vowel, e.g., 3sg/3pl <i>sprîjină</i> “support,” <i>suferă</i> “suffer” (infin. <i>sprîjini/suferi</i>)</p> <p>2. Syncretic form ends in <math>-e</math> (theme vowels <math>-a, -i</math>): a. Non-monosyllabic stem ending in palatal glide plus <math>-a</math> (<math>-ia</math>), e.g., 3sg/3pl <i>taie</i> “cut” (infin. <i>tăia</i>) b. Stems ending in a two-vowel sequence, the first of which is non-front, e.g., 3sg/3pl <i>sue</i> “climb” (infin. <i>sui</i>)</p>
1sg/2sg present	<p>Syncretic form ends in palatal glide (<math>-i</math>): Non-monosyllabic stem ending in palatal glide plus <math>-a</math> (<math>-ia</math>), e.g., 1sg/2sg <i>tai</i> “cut” (infin. <i>tăia</i>)</p>
1sg/1pl imperfect	All verbs
3sg/3pl subjunctive	All verbs

Table 3: *Syncretic types in Romanian conjugation*

For example, several monosyllabic stems, whose only vowel is the stem-final theme (e.g., *da* “give,” *sta* “stay,” *bea* “drink,” *vrea* “want,” *ști* “know,” etc.) all share the same syncretic pattern of 1sg/3pl, realized by the glide [u] (spelled *u*, e.g., 1sg/3pl *dău, stău, beău, vreau, știu*). The phonological reason is that the underlying  $-u$ , found in both the 1sg and 3pl, is changed to the glide when immediately following a stressed vowel, which is inevitably the case in monosyllabic stems. Therefore, the conditioning factor for this syncretic type has nothing to do with the traditional conjugation classes, but is directly related to the phonological feature of monosyllabicity. Thus, even a small sample of verbs should be sufficient to convince the observer of the direct causal role of stem phonology in determining the resulting syncretism.

Likewise, the syncretic pattern of 3sg/3pl is manifested by a final mid vowel: *-ă* (phonetically, [ə]) or *-e*. As in the case of 1sg/3pl syncretism, a wide variety of conjugational types can manifest this syncretism, all of which can be defined on the basis of the specific phonological properties of the stem. Thus, we see that non-monosyllabic stems with a basic non-front theme-vowel (*-a* or *-î*, phonetically [i]) regularly have 3sg/3pl syncretism, but stems with basic front theme vowels only have this type of syncretism in narrowly defined conditions, such as when the front theme vowel is immediately preceded by a back vowel (e.g., infinitive *sui* “climb,” 3sg/3pl *suie*) or is not contiguous to the present-tense stress (e.g., infinitive *sprijini* “support,” 3sg/3pl *sprijină*).

Table 3 also displays the difference between the phonologically conditioned types of syncretism found in the present tense and the non-phonological types found outside the present tense. For example, in the imperfect, first person is syncretic for number and in the subjunctive third person has number syncretism. However, these syncretisms are not based on phonology and apply to all verbs. By definition, if a syncretic type pervades an entire grammatical category, such as the first-person syncretism of the Romanian imperfect, there will be no phonological subcategories that condition the presence or absence of the syncretism. Therefore, it is obvious that the number syncretisms found in both the imperfect and subjunctive are qualitatively different from those of the present tense, where phonological conditioning is the rule.

## 2. *A new segmentation and rule system*

### 2.1 *Segmentation of present-tense and imperfect desinences*

In this section, a new segmentation of Romanian verbs is proposed, with a unified set of grammatical endings. The variable realizations and syncretisms among the various verb stems are not caused by the fact that these stems lexically belong to a particular conjugation, but by the phonological interaction of the stem-final (theme) vowel and the ending. Each grammatical ending of the conjugational system is really a complex of three morphemes, representing the grammatical categories of tense, number, and person, in that order. Conjugation can be defined as the cyclical combination of the stem-final segment with each of the three components of the grammatical ending. The rules are morphophonemic, rather than strictly phonological, in that many of them only occur within the process of conjugation. There is no absolute neutralization of any of the posited grammatical morphemes. Each of the assumed morphemes surfaces in its basic form in at least some environments. In cases where certain morphemes (e.g., word-final *-u*) do not surface, this

occurs due to general phonological rules that exclude such occurrences.<sup>4</sup> My notation will show the basic stem followed by the '+' symbol, for example, *invita+*; components of the grammatical desinence are separated by hyphens. Zero morphemes are rendered with the symbol 'Ø,' and word-final position is indicated with the symbol '#.' Each zero morpheme is opposed by at least one instance of a non-zero.

	Tense	Number	Person		Tense	Number	Person
1sg	Ø	Ø	-u	1pl	Ø	-m	-u
2sg	Ø	Ø	-i	2pl	Ø	-t	-i
3sg	Ø	Ø	Ø	3pl	Ø	-u	Ø

Table 4: *System of present-tense endings*

	Tense	Number	Person		Tense	Number	Person
1sg	-á	-m	-u	1pl	-á	-m	-u
2sg	-á	Ø	-i	2pl	-á	-t	-i
3sg	-á	Ø	Ø	3pl	-á	-u	Ø

Table 5: *System of imperfect-tense endings*

The system of present- and imperfect-tense endings can be found in Tables 4 and 5. The middle morpheme position, that of number, significantly differs from the others, in that there is no basic phonological representation for the plural, although the singular can be treated as a consistent zero. In any case, it is assumed that some prior mechanism will make these morphemes available for conjugation in the way they are represented in Tables 4 and 5. The present-tense morpheme is a zero, and is opposed to the stressed vowel morpheme *-á* of the imperfect.<sup>5</sup> The *-á* does, in fact, surface in all forms of the imperfect. The process of commutation suggests a zero for the present, since certain

<sup>4</sup> One might question the assumption of a final *-u* for first person in the 1pl form, since it can never surface in this form for phonological reasons, that is, it can only be preceded by nasal *m*, while a preceding [consonant + liquid] cluster is required for final *-u* to surface. However, both the parallelism with the 1sg and the possibility of an underlying final *-u* argue in favor of its underlying existence.

<sup>5</sup> I posit stressed *-ú* for the simple perfect. Due to space considerations and the fact that there is no syncretism in the simple perfect, the topic is not discussed in this chapter.



imperfect desinences are longer than the corresponding present-tense endings by one segment (e.g., *bateam* “beat<sub>1SG/PL,IMPF</sub>” vs. *batem* “beat<sub>1PL,PRESENT</sub>”). The consonantal desinences, which represent number in the first and second persons, regularly surface as such, except for the fact that 2pl basic *-t* is palatalized to *-tʃ*, due to its position before a word-final *-i*. In fact, unstressed word-final *-u* and *-i* are subject to the most variation. They survive intact only when preceded by a [consonant + liquid] cluster (e.g., *aflu*, *aflʃi*); when preceded by a single consonant or another type of consonant cluster, the *-u* is deleted (e.g., 1sg *invit*), while the *-i* is reduced to a nonsyllabic glide (*i*), which often palatalizes the preceding consonant (e.g., 2sg *invʃiti*).

## 2.2 Rules for vowel → mid and vowel deletion

In the imperfect tense, first-person syncretism is not phonologically conditioned; I assume that it is grammatically assigned and appears as the input to the phonological rules, as shown in the identical first-person slots of Table 5. By contrast, the syncretic cells of the present tense come into existence due to the application of morphophonemic rules. Two specific types of morphophonemic rules are of particular importance for the generation of present-tense forms; their function will be the next major topic of discussion. Theme vowels all surface as mid at certain points in the present tense, even if they are high or low vowels in the basic stem. This suggests that there is a rule that lowers high vowels to mid and raises low vowels to mid. This is most obvious in the 3sg form, where the corresponding infinitives have high and low stressed vowels (cf. *invítá*, *sāri*), yet these vowels all surface in the present as unstressed vowels that have changed to mid (cf. 3sg *invítă*, *sáre*). Therefore, the first major sort of rule will require non-mid vowels to change to mid in particular environments. When the basic (unstressed) theme vowels totally fail to surface in certain present-tense forms (such as 1sg), even as changed to mid, we can observe that the following underlying desinential vowel surfaces instead, leading to the conclusion that a sequence of two underlying unstressed vowels experiences the deletion of one of these vowels. Since high and low stressed vowels (found in monosyllabic stems) do not undergo deletion in the 1sg, but instead cause the *-u* to become a glide (cf. *ʃtúu*, *dáuu*), I assume that the deleted theme vowels experienced deletion before *-u* because they had first become unstressed mid vowels. In the 1sg, the vowel that survives the deletion (final *-u#*) is itself subject to later deletion, unless a [consonant + liquid] sequence precedes it (e.g., *invit*, but *sufllu*). Therefore, the surface forms do not obviously demonstrate that a vowel sequence and a deletion have occurred. Yet, if we start from basic stems *invita+* and *suflla+*, deletion in both cases is the only

way to generate *invit* and *sufllu*. In the 1sg form, all of the verb classes show a similar type of deletion. However, verbs are dramatically different in the 3pl and that is the key to solving the riddle of how syncretism is generated.

Since front-vowel themes (i.e., *-i* and *-e*) experience the same deletion in 3pl and 1sg, those two forms are syncretic. The results of the commutation process (shown in Table 4) indicate that the underlying grammatical endings of the 1sg and 3pl differ in that 1sg has an *-u* desinence in the person slot (i.e., in word-final position), while 3pl has an *-u* desinence of number, which occupies the desinential medial slot, not directly on the word-final boundary. Since front-vowel themes (*-e*, *-i*) both experience the identical deletion in these two instances (1sg and 3pl, e.g., *bat*, *sar*), I assume that these themes undergo the same rules for being deleted before *-u* in both positions. However, the back-vowel themes (both *-a* and *-î*) experience deletion only in the 1sg (e.g., *invit*, *cobor*), but not in the 3pl (*invită*, *coboară*), where they surface as nondeleted mid vowels. This indicates that the high front theme vowel (*-î*) generalizes mid-vowel height before any unstressed desinential vowel, regardless of whether it is in word-final position, but that the high and low non-front theme vowels (*-a*, *-î*) do not undergo the change to mid-vowel height before a nonfinal *-u* (such as in the 3pl), but do so only at the point in the cycle when they reach the word-final boundary.

Therefore, due to rule ordering, if a basic *-a* theme has not yet been raised to mid, it will not be deleted before a following *-u* desinence; rather, the *-u* desinence itself will be deleted and only later will the *-a* reach the word-final position and be raised to mid, surfacing as *-ă* (cf. 3pl *invită*). If, on the other hand, an *-i* theme gets lowered to mid even before being combined with desinential medial *-u*, the sequence *-e-u-* will ensue and the first vowel will be subject to deletion, first yielding a form such as *bat-u#*, which later loses the final *-u* and surfaces as 3pl *bat*, syncretic with 1sg. Thus, at each cycle of conjugation (stem + tense, stem + number, and stem + person), there will be rules for changing theme vowels to mid-vowel height and then for deleting vowels in the sequence of theme vowel plus vocalic desinence, as follows:

(i) Any unstressed front-vowel theme (regardless of frontness or backness) becomes mid before an unstressed desinential vowel. This rule can take the form (1), where the '+' symbol refers to the stem-desinence boundary:

- (1) V  
 [-back]            [-high]            V  
 [-stress] → [-low] / \_\_\_ + [-stress]

(ii) An unstressed theme vowel becomes mid before a word-final vowel or zero desinence. This can be represented as in (2):

- (2) V                    [-high]  
 [-stress] → [-low] / \_\_\_ + (V) #

(iii) An unstressed mid vowel is deleted when it precedes an unstressed desinential vowel; an unstressed high desinential vowel is deleted when it is preceded by a non-mid (high or low) vowel. In rule form, this will be as in (3) through (5):

- (3) V  
 [-high]  
 [-low]                    V  
 [-stress] → Ø / \_\_\_ + [-stress]

- (4) V                    V  
 [+high]                [+high]  
 [-stress] → Ø / [-stress] + \_\_\_

- (5) V                    V  
 [+high]                [+low]  
 [-stress] → Ø / [-stress] + \_\_\_

Notice that all of these rules apply to sequences of unstressed vowels. The rules for vowel→mid and vowel deletion do not operate as such when the verb stem is monosyllabic and the theme is stressed, which is why verbs such as *da* and *sta* are often treated as irregular.<sup>6</sup> In this case, deletion rule (5) does not apply. Instead of being deleted, the *-u* is changed to the corresponding glide [u̯]. This explains why the 3pl of first-conjugation *da* is *dau* (with final [u̯]), but the 3pl of *invita* is *invitǎ*, with deleted *-u*.

<sup>6</sup> Present-tense stress assignment is by default in the verbs under discussion. The last preconsonantal vowel gets the default stress, unless there is no such vowel, which explains why a polysyllabic stem (e.g., 3pl *invitǎ*) has pretheme stress, but a monosyllabic stem (e.g., 3pl *dau*) has a stressed theme vowel (see Feldstein 1994-1995:239-245).

### 2.3 Summary of rules applied to 3pl forms

The cyclical operations in the critical present 3pl form, at the morphemes of tense, number, and person, for the basic input *invita+Ø-u-Ø#*, *bate+Ø-u-Ø#*, and *sari+Ø-u-Ø#* can be briefly summarized as follows:

(i) In the present tense, the first desinential position (tense morpheme) is a zero, so no vowel combinations occur.

(ii) Theme vowel → mid: When the second position (number) is reached, the unstressed high front-vowel theme becomes mid, that is,  $i \rightarrow e$ ; *sar-i+u+Ø#* → *sar-e+u-Ø#*. Back-vowel themes are exempt from this rule, since they only change to mid when the following desinence is word final. In this case, the notion of ‘word final’ is abstract, since a zero occupies the final position.

(iii) Vowel deletion: An unstressed mid vowel is deleted before a high vowel: *sar-e+u-Ø#* → *sar-é+u-Ø#* → *sar+u-Ø#*. When a low vowel is followed by a high vowel, it is the high vowel that gets deleted: *invita+u-Ø#* → *invita+ú-Ø#* → *invita+Ø#*.

(iv) Next, at the third(-person) morpheme, the theme vowel → mid rule calls for all unstressed theme vowels to become mid when preceding a word-final desinence: *invita+Ø#* → *invitǎ-Ø#* → *invitǎ*.

(v) Eventually, a postlexical rule calls for the deletion of final *-u*, unless it is preceded by an [obstruent + liquid] cluster: *sar+u-Ø#* → *sar+u#* → *sar*.

To recapitulate, both 1sg and 3pl both contain an *-u* morpheme, but differ in that the 1sg *-u* is in the third slot, while the 3pl *-u* is in the second slot. Theme *-i* becomes mid at the second slot and theme *-e* is mid to begin with, but theme *-a* becomes mid only at the third slot. This accounts for the different syncretisms of the so-called I versus the other conjugations of Romanian.

In order to further clarify the process of how the rules of theme → mid and vowel deletion derive the present-tense forms, Tables 6 through 8 illustrate the step-by-step process for three different theme vowels (*-a*, *-e*, *-i*) in the 1sg, 3sg, and 3pl forms. The other non-front-vowel theme (*-î*) behaves like *-a*.

1. Basic input	invita+Ø-Ø-u#	bate+Ø-Ø-u#	sari+Ø-Ø-u# <sup>7</sup>	
2. Tense cycle	invita+Ø-u#	bate+Ø-u#	sari+Ø-u#	Elimination of zero tense morpheme
3. Number cycle	invita+u#	bate+u#	sari+u#	Elimination of zero number morpheme
4. Person cycle	(a) invitə+u#	-----	sare+u#	Unstressed theme vowel → mid preceding word-final desinence
	(b) invitə+u #	baté+u#	saré+u#	Deletion of unstressed mid [e/ə] before desinential vowel
5. Postlexical:	invit	bat	sar	Loss of -u# unless blocked due to preceding [consonant + liquid] sequence
Orthographic form	<i>invit</i>	<i>bat</i>	<i>sar</i>	

Table 6: Derivation of the verbs *invita*, *bate*, and *sāri* in 1sg (vacuous rule applications are shown with dotted lines)

1. Basic input	invita+Ø-Ø-Ø-#	bate+Ø-Ø-Ø-#	sari+Ø-Ø-Ø-#	
2. Tense cycle	invita+Ø-Ø#	bate+Ø-Ø#	sari+Ø-Ø#	Elimination of zero tense morpheme
3. Number cycle	invita+Ø#	bate+Ø#	sari+Ø#	Elimination of zero number morpheme
4. Person cycle	(a) invitə+Ø#	-----	sare+Ø#	Unstressed theme vowel→mid preceding word-final desinence
	(b) invitə#	bate#	sare#	Elimination of zero person morpheme
Orthographic form	<i>invitǎ</i>	<i>bate</i>	<i>sare</i>	

Table 7: Derivation of the verbs *invita*, *bate*, and *sāri* in 3sg

<sup>7</sup> I recognize the basic stem as *sari+*. The root vowel of the infinitive appears as the mid vowel [ə] (*ǎ*, i.e., *sāri*), due to the fact that it is unstressed. When stressed, the vowel is *a*, as in 1sg/3pl *sar*, and 3sg *sare*. However, the rules for raising unstressed vowels do not directly bear on the issues of syncretism that are the focus of this chapter.

1. Basic input	invita+Ø-u-Ø#	bate+Ø-u-Ø#	sari+Ø-u-Ø#	
2. Tense cycle	invita+u-Ø#	bate+u-Ø#	sari+u-Ø#	Elimination of zero tense morpheme
3. Number cycle	(a) -----	-----	sare+u-Ø#	Unstressed front vowel → mid before vocalic desinence
	(b) -----	baté+u-Ø#	saré+u-Ø#	Deletion of unstressed mid before vocalic desinence
	(c) invita+ú-Ø#	-----	-----	Deletion of vocalic desinence after (non-mid) theme vowel
4. Person cycle	(a) invitə+Ø#	-----	-----	Unstressed theme vowel → mid preceding word-final desinence
	(b) invitə#	bat+u#	sar+u#	Elimination of zero person morpheme
5. Postlexical	-----	bat	sar	Loss of -u# unless blocked due to preceding [consonant + liquid] sequence
Orthographic form	invită	bat	sar	

Table 8: Derivation of the verbs *invita*, *bate*, and *sări* in 3pl

**3. Romanian syncretism as depicted by Stump (2001) and Bobaljik (2002)**

Having presented a system that treats phonological syncretism as underlyingly nonsyncretic, I return to examining the questionable statements of Stump (2001) and Bobaljik (2002) on the topic of Romanian syncretism.

**3.1 Present-tense 1sg/3pl syncretism**

Stump (2001:213) introduces this type of syncretism by stating that, “in Romanian, verbs belonging to any but the first conjugation have present indicative paradigms in which the 1sg form is identical to the 3pl form.” However, as noted earlier, first-conjugation verbs with monosyllabic stems can indeed have this type of syncretism, for example, 1sg/3pl *dau/stau*. It is only necessary that they fulfill a specific phonological condition: First-conjugation verbs with this syncretism must have a theme vowel that is consistently stressed in the present tense. Therefore, the issue is not an abstract grammatical assignment called ‘first conjugation,’ but the phonological fact of whether a theme vowel -a is stressed or unstressed in the 1sg and 3pl forms in question. In other words, the reason is phonological, rather than morphological. I understand the 3pl to contain an underlying -u desinence that is retained as a

glide after a stressed theme vowel *-a*, but that gets deleted after an unstressed *-a* theme (cf. first-conjugation 3pl with stressed theme vowel *dau, stau* vs. first-conjugation 3pl with unstressed theme vowel *invită, suflă*). In spite of the possible surface deletion after the unstressed theme vowel *-a*, I regard the basic 3pl desinence to be *-u*. Stump attempts to apply a rule of referral to the 1sg/3pl syncretism, declaring the 1sg *-u* to be the independent entity, with the 3pl as dependent. This ignores the important fact that, in every instance where the 3pl lacks the *-u*, this absence is attributable to a phonological deletion, such as that which occurs after an unstressed theme vowel *-a*. Bobaljik (2002:66) criticizes Stump's declaration of 1sg as having a more inherent *-u* ending than the 3pl when they are syncretic, on the basis of the fact that the imperfect tense has a constant *-u* in the 3pl, but he does not point out the difference between phonologically conditioned syncretisms and those that have no such phonological restriction. In fact, both Stump and Bobaljik appear unsure about the precise nature of the *-u* desinence, since it can appear in both 1sg and 3pl, and is subject to a variety of phonological restrictions in both paradigmatic slots. Ultimately, Bobaljik concludes that the *-u* is simply a default form, due to "this rather scattered distribution of the *-u* suffix" (2002:66). The solution I proposed in section 2 claims that both 1sg and 3pl contain an *-u* component, but that the two forms differ in that 1sg has this desinence directly on the word-final boundary, while 3pl has it preceding a zero element. In the case of 1sg, I would claim that the *-u* signals first person, but that it represents plural number in the 3pl.

Table 2 shows that the present paradigms of *bate* and *umple* can have 1sg/3pl syncretism, with surface realizations of both  $\emptyset$  and *-u*, respectively. While Stump (2001) does not even mention the option of zero in his discussion of this type, Bobaljik (2002) states that it is a special property of the 1sg/3pl syncretism, calling it 'metaparadigmatic.' However, such multiple realizations are quite common among the phonologically conditioned syncretisms of Romanian, that is, they are not confined to just one type. For example, the 3sg/3pl type of syncretism, normally associated with the first conjugation, also has two realizations, depending on whether a semivowel [ɨ] precedes the theme-vowel, since the usual schwa fronts to *e* when preceded by [ɨ] (e.g., *tăia* is realized as *tăie* in both 3sg and 3pl, as shown in Tables 2 and 3). Therefore, Bobaljik's so-called metaparadigmatic option occurs not only for 1sg/3pl, but also for 3sg/3pl.

### 3.2 *Present-tense 3sg/3pl syncretism*

As mentioned earlier, Stump (2001) and Bobaljik (2002) identify 1sg/3pl syncretism with the traditional non-first conjugation types. Following this pattern, they identify the other major type of present-tense syncretism (3sg/3pl) with the first conjugation. Yet, there are several verb classes that do not have the first-conjugation theme vowel *-a*, yet do have this type of syncretism (see the 3sg/3pl section of Table 3). They are not just inexplicable irregulars, but clear cases of phonological subtypes, whose environments create the conditions for this particular type of syncretism. In the first place, even theme vowel *-a* has a phonological restriction on its use with 3sg/3pl syncretism: This theme vowel cannot belong to a monosyllabic stem and have constant stress in the present; otherwise, it has 1sg/3pl syncretism, for example, *da/sta*. Theme vowels outside the first conjugation (i.e., other than *-a*) can also regularly have 3sg/3pl syncretism, as long as they are either back vowels (such as *-î*, e.g., *coborî* “descend”), front vowels that are immediately preceded by a back vowel (e.g., *sui* “climb”), or when there is an underlying stress mark two syllables to the left of the theme<sup>8</sup> (*sprijini* “support,” *suferi* “suffer”).

Stump (2001:213-214) states that 3sg/3pl syncretism presents “a rule of *-ă* suffixation which expresses third person but is insensitive to differences of number.” This raises objections from the perspective of the morphophonemic system I introduced earlier. In the first place, I regard the *-ă* ([ə]) simply as a raised theme vowel, followed by a zero ending, rather than as a present-tense desinence on its own. Second, when the very same 3sg *-ă* happens to be the final vowel of a stressed monosyllabic stem (stressed *dă*) there is no 3sg/3pl syncretism (3sg *dă* vs. 3pl *dău*). These are just the sorts of grammatical conclusions that should not be made with such a broad brush when they are restricted to particular phonological environments.

### 3.3 *Present-tense 1sg/2sg syncretism*

In addition to the other and more widespread syncretisms of the present tense, there is also the minor and restricted 1sg/2sg syncretism, which is not mentioned by either Stump (2001) or Bobaljik (2002). It is phonologically conditioned when verbs have a [j] consonant that precedes the theme vowel, for example, 1sg/2sg *tai* “cut,” shown in Table 2. This syncretism never appears alone and must co-occur with 3sg/3pl syncretism, producing a double syncretism within a single present-tense paradigm (e.g., 1sg/2sg *tai*, plus

<sup>8</sup> For more details on the relationship of this stress pattern to Romanian conjugation, see Feldstein (1994-1995:240).



3sg/3pl *taie*). It does not combine with the 1sg/3pl type; if it did, it would produce a situation of syncretism across three cells. This may indicate that a single tense paradigm can tolerate two syncretic pairs on the surface, but not a syncretic triplet.

### 3.4 *Imperfect 1sg/1pl and subjunctive 3sg/3pl*

The structure of the imperfect-tense syncretism lacks any phonological conditioning. Therefore, it occurs throughout all verb types and there is no instance in which Romanian distinguishes 1sg and 1pl in the imperfect. In all of the present-tense instances looked at hitherto, the phonological conditioning goes along with the fact that each type is found with a delimited number of verb stems, which can be defined phonologically.

The syncretic situation of the Romanian subjunctive was not mentioned in the discussions of either Stump (2001) or Bobaljik (2002). It has unique syncretic properties in the Romanian system of verbal syncretism, perhaps due to the fact that it is the only instance of a mood opposition. In the first two persons, there is syncretism between the present indicative and subjunctive, which could be treated as impoverishment, based on the absence of a mood opposition outside the third person. In the third person, the mood opposition is present, but the subjunctive 3sg is the same as the 3pl in all verbs, regardless of stem phonology. Therefore, on the basis of no phonological conditioning and invariant number syncretism, the imperfect first-person and subjunctive third-person syncretisms are structurally similar. Both stand in stark contrast to the present, in which there is no invariant number syncretism. The only invariant syncretic cell (i.e., having obligatory syncretism) in the present tense is that of the 3pl. Since the 3pl present is always syncretic with either 1sg or 3sg, one can say that the number syncretism is also an invariant of the 3pl present, although it can take the form of combined person/number syncretism in the case of 1sg/3pl, or pure number syncretism, in the case of 3sg/3pl. The common denominator for present, imperfect, and subjunctive is that each paradigmatic instance contains two forms that suppress number, while expressing either present, imperfect, or subjunctive meanings.

## 4. *Do patterns of surface syncretism contribute anything?*

The proposed system of segmentation (Tables 4 and 5) and rules (Tables 6-8) posits basic shapes for the grammatical morphemes of tense, number, and person, and establishes rules that can derive the seemingly capricious patterns of Romanian syncretism. However, this approach does not necessarily exclude an analysis of what the surface phenomena reveal, which may show another

side of the functioning of the syncretic system.<sup>9</sup> In other words, even though the syncretism is not underlying, there may be an inner logic to the rules, which leads to surface patterning. The main lines of surface syncretism can be summarized as follows. The 3pl is obligatorily syncretic in the present, but all three persons of the singular (1sg/2sg/3sg) are optionally syncretic, depending on the phonology of the particular verb stem. One of two syncretisms must occur in every present-tense paradigm: either 1sg/3pl or 3sg/3pl. Therefore, since only first- and third-person forms are found within the set of obligatory types, second person is excluded when there are only two syncretic present-tense cells. However, if a second syncretism occurs (normal for stems with a palatal glide preceding the theme vowel, e.g., *tăia*), one syncretic pair must pertain to the third person (3sg/3pl), while the other (1sg/2sg) excludes third person. The optional second syncretic pair (1sg/2sg) is thus notable in that it is the precise opposite of the obligatory form that must always be syncretic in the present (3pl), since this optional type is defined as non-third person and non-plural in number. In other words, there is a negative correlation between the two syncretic patterns (3sg/3pl and 1sg/2sg) that can co-occur in the same paradigm.

To summarize, all of the present-tense syncretisms can make either positive or negative reference to second or third persons, but never first person, as follows, based on the common denominator of each of the syncretic types:

- (i) 3sg/3pl (*invită*) has the common denominator of third person.
- (ii) 1sg/3pl (*bat*) has the common denominator of non-second person.
- (iii) 1sg/2sg (*tai*) has the common denominator non-third person.

Conversely, the past-tense syncretism (of the imperfect) is diametrically opposite to the lack of first-person reference in the present tense. Past-tense syncretism is obligatorily defined as first person (1sg/1pl, e.g., *invitam*), precisely the type that is excluded in the present. Therefore, we can state that none of the varieties of present-tense syncretism ever make explicit reference to first person, while past-tense syncretism can only make explicit reference to the first person. As noted earlier, present-tense syncretism comes about as a

---

<sup>9</sup> An anonymous reviewer wondered how I can first propose underlying forms to derive syncretic forms and then say that the surface is also of interest. I do claim that both the derivation and the surface are interesting structures from differing points of view, much as one might construct both a grammar for the speaker and the listener, or study both morphophonemic and phonetic patterns in a language. One might recall that Roman Jakobson, who wrote the pioneering generative study of the Russian verb (Jakobson 1948), also devoted many studies to the subject of the surface patterning of syncretisms (e.g., Jakobson 1958).

result of phonological rules, while in the imperfect, the syncretism appears to be imposed by the morphological component, prior to the functioning of the phonological rules. Curiously, if the 1sg/1pl imperfect syncretism had not been assigned by the grammar, there would have been an invariant phonological syncretism in the imperfect, of the type 1sg/3pl, since the imperfect-tense morpheme (stressed *á*) would combine with either the third-position *-u* of the 1sg or the second-position *-u* of the 3pl in the same way; that is, *invitau* would not only be the 3pl form (as it actually is), but would have been a 1sg/3pl syncretic form, if it were not for the nonphonological imposition of the 1sg/1pl syncretism. However, if 1sg/3pl syncretism were to exist in the imperfect, there would then be no special oppositional role of the first person, as excluded from the present but obligatory in the past. One can only speculate about the conspiratorial role of the grammatical system in its imposition of non-phonological first-person syncretism in the imperfective. Stump (2001) and Bobaljik (2002) differ in their explanations of the unusual fact that imperfect syncretism is realized by the marked plural, rather than the unmarked singular (Bobaljik 2002:65). My system of segmentation suggests the view that the language opted for a marked and nonphonological choice (1sg/1pl), due to the fact that the phonology would have yielded an imperfect-tense syncretism not strictly of the first-person type, which would have gone against the overall grammatical strategy of syncretism (i.e., absence of present-tense reference to first person, but imperfect reference only to first person).

The syncretism of the subjunctive mood manifests structural similarities to both the present and the imperfect. It is like the present in its syncretism of the third person (3sg/3pl, e.g., *invite*, *bată*); yet, it has the pattern of the imperfect in its single, constant number syncretism within a single grammatical person (third), the only number in which the subjunctive is opposed to the present tense.

## 5. Conclusion

This chapter has attempted to demonstrate some of the complexities and systematic properties of Romanian conjugational syncretism, particularly when it is phonologically conditioned in the present tense. Two recent linguistic discussions of Romanian syncretism were shown to contain inaccuracies, due to the fact that they follow the traditional view of linking syncretism to conjugational class. It was demonstrated that one cannot accurately deal with syncretism in terms of the traditional conjugational types. Phonologically conditioned syncretism, as found in the Romanian present tense, is structurally very different from the nonphonological type, as found in the imperfect. As an

alternative to the traditional conjugational types, a new morphophonemic system was proposed in which verb stems, including the theme vowel, all cyclically combine with a tripartite desinential complex. As shown, this system is capable of generating the correct syncretic forms and can help to explain how syncretism arises in the structure of the language. Thus, this chapter has indicated the inadequacies of previous analyses and suggested a systematic alternative for the analysis of Romanian syncretism. Finally, a surface-oriented analysis of syncretism was explored, which might serve to complement the analysis of underlying forms. Both the underlying morphophonemic forms and the surface forms constitute a complex structure that has yet to be fully understood in all of its ramifications.

#### REFERENCES

- Bobaljik, Jonathan David. 2002. "Syncretism Without Paradigms: Remarks on Williams 1981, 1994". *Yearbook of Morphology 2001* ed. by Geert Booij & Jaap van Marle, 53-85. Dordrecht: Kluwer.
- Feldstein, Ronald F. 1994-1995. "Accentul verbului în limba română". *Dacoromania* 1:1-2.75-96.
- Graur, Al., ed. 1966. *Gramatica limbii române*, vol. I. Bucharest: Editura Academiei Republicii Socialiste România.
- Jakobson, Roman. 1948. "Russian Conjugation". *Word* 4:5.155-167.
- . 1958. "Morfologičeskie nabljudenija nad slavjanskim sklonenijem". *American Contributions to the Fourth International Congress of Slavists*, 127-156. The Hague: Mouton.
- Stump, Gregory T. 2001. *Inflectional Morphology: A theory of paradigm structure*. Cambridge: Cambridge University Press.
- Zwicky, Arnold M. 1985. "How to Describe Inflection." *Proceedings of the eleventh annual meeting of the Berkeley Linguistics Society: 16-18 February, 1985* ed. by Mary Niepokuj, 372-286. Berkeley, CA: Berkeley Linguistics Society.



**SLUICING IN ROMANIAN\***  
**A TYPOLOGICAL STUDY**

FREDERICK HOYT & ALEXANDRA TEODORESCU  
*University of Texas at Austin*

**0. Introduction**

‘Sluicing’ (Ross 1967) refers to English constructions in which the clausal subconstituent of a question is elided, leaving a ‘floating’ *wh* phrase (or ‘remnant’). For example, in (1) the remnant is *who* and the ellipse is understood as *she is marrying*.

- (1) Arabelle is marrying someone you know! Guess who!

Similar constructions exist in Romanian (2) and Japanese (3):

- (2) *Cineva mi-a mâncat prăjiturile, dar nu știu cine.*  
someone CL.1S-PAST.3S eaten<sub>PAST.3S</sub> cookies-the but not know<sub>1S</sub> who  
“Someone ate my cookies but I don’t know who.”
- (3) *Bill-ga nanika-o nusunda rasiu kedo, watashi-wa nani-o (da) ka siranai.*  
Bill-NOM something-ACC stole seembut I-TOP what-ACC is Q  
know-not  
“It seems that Bill stole something, but I don’t know what.”

English sluicing has been analyzed as IP-ellipsis (Lobeck 1995; Merchant 1998, 2001), as in (4):

- (4) guess [<sub>CP</sub> who<sub>i</sub> [<sub>IP</sub> ~~she is marrying~~ t<sub>i</sub>]].

Shimoyama (1995), Merchant (1998), and Hiraiwa and Ishihara (2002) argue that while Japanese examples like (5) resemble English sluicing, they actually

---

\* Thanks to: Bernhard Schwarz, Rajesh Bhatt, Danny Fox, Lisa Green, Jason Merchant, Junko Shimoyama, Steve Wechsler, and to the anonymous reviewers for discussion and comments on this work; to Masa Deguchi, Hitoshi Hirioishi, Makiko Irie, and Tomoko Sakuma for their help with Japanese data; to Dan Tecuci for his assistance with Romanian; and to Kathryn Tippetts for hosting us at LSRL 33, and for generally making our presentation possible.

are ellipsis of the CP-constituent in a cleft construction, rather than of the IP-node of a matrix clause:

- (5) *watashi-wa* [<sub>CP</sub> ~~*Bill-ga nusunda no*~~] *nani-o (da) ka siranai*.  
 I-TOP Bill-NOM stole COMP what-ACC is Q know-not  
 “I don’t know what [it is that Bill stole].”

Romanian and Japanese sluicing are similar in ways that exclude English, at least at the descriptive level: They allow multiple remnants, they allow non-*wh* remnants, and they allow overt complementizers in the sluice. Given these similarities, it is natural to ask if Romanian and Japanese sluicing might have similar analyses. The topic of this chapter is whether Romanian sluices like (2) pattern with English or Japanese sluices in terms of their structure; in other words, whether (6a) or (6b) is a more appropriate structural analysis for the Romanian sluice in (2).

- (6) a. *dar nu știu cine<sub>i</sub> [<sub>st<sub>i</sub></sub> ~~*mi a mâncat prăjiturile*~~].*  
 but not know<sub>IS</sub> who CL.IS PAST.3S eaten cookies-the  
 “(Someone ate my cookies) but I don’t know who [ate them].”  
 b. *dar nu știu cine<sub>i</sub> [<sub>S t<sub>i</sub></sub> [<sub>CP OP<sub>i</sub></sub> ~~*mi a mâncat prăjiturile*~~]].*  
 but not know<sub>IS</sub> who CL.IS PAST.3S eaten cookies-the  
 “(Someone ate my cookies) but I don’t know who [it is who ate them].”

We find that, despite the superficial similarities between Romanian and Japanese sluicing, an S-ellipsis analysis of Romanian sluicing as in (6a) is to be preferred. We show that the similarities between Romanian and Japanese sluicing are epiphenomenal and follow from independent syntactic properties of the two languages. An S-ellipsis analysis immediately accounts for the key properties of Romanian sluicing. The differences between Romanian and English sluicing follow from the presence of a richer structure in the left periphery of embedded clauses in Romanian.

A terminological note is in order. We use the term ‘sluicing’ exclusively as a cross-linguistic descriptive label for ellipsis of subconstituents of an embedded question. This has subcategories such as English Sluicing, Romanian Sluicing, and Japanese Sluicing. As we use it, sluicing implies nothing about the structural analyses of the data. Instead, analyses are referred to as different kinds of ‘ellipsis’: English and Romanian sluices are both subclasses of S-ellipsis, while Japanese sluicing is a class of CP-ellipsis.

The chapter is organized as follows: In section 1, we discuss the superficial similarities between Romanian and Japanese sluicing, and how these similarities contrast with English sluicing. In section 2, we discuss the crucial

similarity between English and Romanian sluicing: island insensitivity. In section 3 we compare possible analyses for Romanian sluicing and conclude that an S-ellipsis analysis is the only option, providing evidence that supports such an analysis. Section 4 concludes the chapter.

### 1. *Similarities between Romanian and Japanese Sluices*

Romanian and Japanese sluices share several properties. These include multiple *wh* remnants (1.1), aggressively non-d-linked *wh* remnants (1.2), overt complementizers (1.3), and a variety of non-*wh* remnants (1.4), like semantically referential, quantificational, and polarity-sensitive remnants (Shimoyama 1995; Merchant 1998). This contrasts with English, which allows only singleton *wh* remnants.<sup>1</sup>

#### 1.1 *Multiple wh remnants*

Romanian and Japanese sluicing both allow multiple remnants (shown in bold) in the sluice (7a-b), while English does not (8a-b):

- (7) a. *Ion a dat **cuiva ceva**, și vreau să știu  
 Ion PAST.3S given someone<sub>DAT</sub> something and want<sub>1S</sub> SUBJ know<sub>1S</sub>  
**cui ce.**  
 whom<sub>DAT</sub> what  
 “John gave something to someone, and I want to know what to whom.”*
- b. *Taro-ga **dareka-ni nanika-o** ageta rasii ga boku-wa **dare-ni**  
 Taro-NOM someone-DAT something-ACC gave heard but I-TOP who-DAT  
**nani-o da ka wakara-nai.**  
 what-ACC is Q know-not  
 “I heard that Taro gave someone something, but I don't know who what.”*
- (8) a. \*John gave **someone something**, and I want to know **who what**.  
 b. ?John gave **something to someone**, but I don't know **what to** whom.

#### 1.2 *Aggressively non-d-linked wh remnants*

Both Romanian and Japanese allow aggressively non-d-linked *wh* words (Pesetsky 1987) as remnants (9a-b), while English does not (10):

- (9) a. *Cineva mi-a ascuns cheile și aș vrea să știu  
 someone CL.1S-PAST.3S hidden keys-the and OPT.1S want SUBJ know<sub>1S</sub>  
 și eu **cine dracu.**  
 even I who devil-the  
 “Someone hid my keys on me, and I'd like to know who-the-hell.”*

<sup>1</sup> Richards (1997) noted that multiple sluicing becomes more acceptable in English if the remnants are PPs and/or nonargumental.



- b. *Minna-ga awateteiru kedo, boku-wa ittai nande ka sirainai.*  
 everyone-NOM panic but I-TOP hell why Q know-not  
 “Everyone is panicking, but I don’t know why the hell.”

(10) a.??Someone ate my sandwich, and I would really like to know who-the-hell.

### 1.3 Overt complementizers

Both Japanese and Romanian tolerate an overt complementizer in the remnant of the sluice. This complementizer can be either interrogative (11) or indicative (12):

- (11) a. *Am aflat că cineva a plecat, dar nu știu dacă Ion.*<sup>2</sup>  
 PAST.1S learned that someone PAST.3S left but not know<sub>1S</sub> if Ion  
 “I found out that someone left, but I don’t know if Bill.”

- b. *John-ga dareka-o kubinisita rasi kedo, boku-wa Bill ka dooka sirainai.*  
 John-NOM someone-ACC fired seembut I-TOP Bill Q whether  
 know-not  
 “It seems that John fired someone, but I don’t know if Bill.”

- (12) a. Dan: *Cine crezi că a câștigat premiul întâi?*  
 who think<sub>1S</sub> that PAST.3S won first-the prize  
 “Who do you think won first prize?”

Alex: *Știam că Anca.*  
 know<sub>PAST1S</sub> that Anca.  
 “I know [that] Anca [did].”

- b. *John-ga dareka-o kubinisita rasi kedo, boku-wa Bill to omou.*  
 John-NOM someone-ACC fired seem but I-TOP Bill that think  
 “It seems that John fired someone, and I think that [it was] Bill.”  
 (Merchant 1998:9)

English does not tolerate overt complementizers in the remnant of a sluice:

- (13) \*One of the foreign students won the department fellowship, and I wonder **whether/if** Joanna.

### 1.4 Non-wh remnants

Both Japanese and Romanian allow a variety of non-*wh* remnants, such as referential NPs (14), adverbs (15), and PPs (16):

<sup>2</sup> An anonymous reviewer points out that there is judgment variation in the acceptability of sluices containing *dacă* “whether/if.” Seven out of ten speakers that we tested found examples (11a), (14a), and (24c) perfectly acceptable while the other three thought they were odd but not ungrammatical. Example (18a) was accepted by all ten speakers.

- (14) a. *Mi s-a spus că cineva s-a întâlnit cu cineva,*  
 Me REFL-PAST.3S tell that someone REFL-PAST.3S met with someone  
*și mă întreb dacă Ion cu Maria.*  
 and me<sub>REFL</sub> wonder if Ion with Maria  
 “I was told that someone met with someone, and I wonder if Ion with Maria.”
- b. *John-ga dareka-o kubinisita rasii kedo, boku-wa Bill-o to*  
 John-NOM someone-ACC fired seems but I-TOP Bill-ACC that  
*omou.*  
 think  
 “It seems that John fired someone, and I think Bill.”
- (15) a. *Carmen vrea să-și ia mașină, și suspectez că [Adv repede].*  
 Carmen wants SUBJ-GEN.F.S take car and suspect<sub>IS</sub> that quickly  
 “Carmen wants to buy herself a car, and I suspect [that] pretty soon.”
- b. *Hanako-wa kuruma-o kaitagatte iru, suguni da to omou.*  
 Hanako-TOP car-ACC buy-want ASP soon is C think  
 “Hanako wants to buy a car, [and] I suspect that soon.”
- (16) a. *Da, am aflat și eu că Ioana a fugit cu cineva, dar*  
 yes PAST.1S learned and I that Ioana PAST.3S elopedwith somebody but  
*n-aș paria că cu Radu.*  
 not-OPT.1S bet that with Radu  
 “Yes, I found out too that Ioana ran off with somebody, but I wouldn't bet that with Radu.”
- b. *Akiko-ga dareka-to kakeochisita to kiita kedo, Taroo-to to-wa*  
 Akiko-NOM someone-with eloped C heard but Taroo-with C-TOP  
*omowanakatta.*  
 not-expected  
 “I heard that Akiko eloped with someone, but I didn't expect Taroo.”

Japanese and Romanian both allow a variety of strong quantificational NPs as remnants, like Romanian *toți* and Japanese *minna* (both “everyone”), and polarity sensitive quantifiers like Romanian *oricine* or Japanese *daredemo* (both “anyone”), as in (17) and (18):

- (17) a. *Da, e adevărat că mulți au votat pentru Iliescu, dar nu*  
 yes is<sub>3S</sub> true that many PAST.3PL voted for Iliescu but not  
*cred că toți...*  
 believe<sub>IS</sub> that everyone  
 “Yes, it is true that many people voted for Iliescu, but I don't believe that everyone.”
- b. *Dareka-ga kono-kuruma-o naoseru to omou kedo, minna-ga to-wa*  
 someone-NOM this-car-ACC can-fix C think but everyone Q-TOP  
*omowanai.*  
 think-not  
 “Someone can fix this car, but I don't know if everyone.”

- (18) a. *Știu că profesorul ajută pe multă lume, dar mă întreb*  
 know<sub>1S</sub> that professor-the help<sub>3S</sub> ACC many people but me<sub>REFL</sub> wonder  
*dacă pe oricine oricând.*  
 if ACC anyone anytime  
 “I know that the professor helps many people, but I wonder if he helps anyone  
 anytime.”
- b. *Dareka-ga kono-kuruma-o naoseru to omou kedo, daredemo ka-wa*  
 someone-NOM this-car-ACC can-fix C think but anyone Q-TOP  
*wakaranai.*  
 know-not  
 “Someone can fix your car, but I don’t know if [just] anybody.”

English, on the other hand, allows only *wh* remnants. Non-*wh* remnants of any category are degraded or unacceptable, as in (19):

- (19) a. ?? I heard that Mary is marrying someone I know; I wonder if **John**.  
 b. ?? Yes, I also heard that Ioana has eloped with somebody, but I wouldn’t bet that  
**with Radu**.  
 c. ?? Hanako wants to buy a car, [and] I suspect that **soon**.  
 d. ?? Luis says that some people from our class cheated on the exam, but I don’t  
 think **everyone**.  
 e. ?? Someone can fix your car, but I don’t think [just] **anybody**.

## 2. *Similarities between Romanian and English sluicing*

Despite these similarities between sluicing in Japanese and Romanian, they differ in one crucial respect, namely that Japanese sluices are island sensitive, while Romanian ones are not. Romanian patterns with English: Both seem to violate island constraints, such as Ross’ (1967) coordinate structure constraint, complex-NP constraint, relative-clause island constraint, sentential subject constraint, and adjunct constraint. That is, the remnant seems to have been extracted across the boundary of a syntactic island.

These island violations are illustrated in (20) for coordinate structure islands, in (21) for complex-NP islands, in (22) for relative-clause islands, and in (23) for adjunct islands:

- (20) a. He invited [<sub>NP</sub> Akiko and someone else ], but I don’t know **who**<sub>i</sub> ( \*he invited  
 [<sub>NP</sub> Akiko and t<sub>j</sub> ] ).  
 b. *Dan a invitat-o* [<sub>NP</sub> *pe Anca și pe încă cineva* ], *dar nu*  
 Dan PAST.3S invited-CL.F.S ACC Anca and ACC other someone but not  
*știu pe cine*<sub>i</sub> ( \* *Dan a invitat-o* [<sub>NP</sub> *pe Anca și t<sub>j</sub>* ] ).  
 know<sub>1S</sub> ACC who Dan PAST.3S invited-CL.F.S ACC Anca and  
 “Dan invited Anca and someone else, but I don’t know who.”
- (21) a. Jerry heard [<sub>NP</sub> a rumor that someone burnt the archive down], but I don’t know  
**who**<sub>i</sub> ( \*Jerry heard [<sub>NP</sub> a rumor that t<sub>j</sub> burnt the archive down ] ).

- b. *Emil a împrăștiat* [<sub>NP</sub> *zvonul că cineva a dat foc arhivei*], *și sunt curioasă cine<sub>i</sub>* (\**Emil a împrăștiat* [<sub>NP</sub> *zvonul arhivei*]<sub>DAT</sub> and *be<sub>IS</sub> curioasă<sub>F,S</sub> who Emil PAST.3S spread rumor-the că t<sub>i</sub> a dat foc arhivei*]).  
that PAST.3S given fire archive<sub>DAT</sub>  
“Emil spread the rumor that someone set the archive on fire, and I wonder who.”
- (22) a. Ana drives [<sub>NP</sub> a car [<sub>CP</sub> that belongs to somebody else]], but I don’t know **who<sub>i</sub>** (\*Ana drives [<sub>NP</sub> a car [<sub>CP</sub> that belongs to t<sub>i</sub>]]).  
b. *Ana conduce* [<sub>NP</sub> *o mașină* [<sub>CP</sub> *care este a altcuiva*]], *dar nu știu a cui<sub>i</sub>*; (\**Ana conduce* [<sub>NP</sub> *o mașină care este t<sub>i</sub>*]).  
GEN who<sub>DAT</sub> Ana drives a car that is GEN somebody-else but not I-know  
“Ana drives a car that is somebody else’s, but I don’t know whose.”
- (23) a. The victim left [<sub>Adjunct</sub> after one of the linguists ], but I don’t know **which** (\*the victim left [<sub>Adjunct</sub> after t<sub>i</sub>]).  
b. *Victima a plecat* [<sub>Adjunct</sub> *după unul dintre lingviști*], *dar nu știu după care* (\**victima a plecat t<sub>i</sub>*).  
victim-the PAST.3S left after one of-the linguists but not know<sub>IS</sub>  
after which victim-the PAST.3S left  
“The victim left after one of the linguists, but I don’t know which.”

Romanian sluices with non-*wh* remnants show the same island insensitivity as those with *wh* remnants, indicating that a similar structure underlies the two classes of examples, as in (24):

- (24) a. *Dan a invitat-o* [<sub>NP</sub> *pe Anca și pe încă cineva*]; *bănui că pe Elena<sub>i</sub>* (\**Dan a invitat-o* [<sub>NP</sub> *pe Anca și t<sub>i</sub>*]).  
Dan PAST.3S invited-CL.F.S ACC Anca and ACC other someone suspect<sub>IS</sub>  
that ACC Elena Dan PAST.3S invited-CL.F.S ACC Anca and  
“Dan invited Anca and someone else; I suspect that Elena.”
- b. *Emil a împrăștiat* [<sub>NP</sub> *zvonul că cineva a dat foc arhivei*], *și eu suspectez că George<sub>i</sub>* (\**Emil a împrăștiat* [<sub>NP</sub> *zvonul că t<sub>i</sub> a dat foc arhivei*]).  
rumor-the that PAST.3S given fire archive<sub>DAT</sub>  
“Emil spread the rumor that someone’s set the archive on fire, and I suspect that George.”
- c. *Ana conduce* [<sub>NP</sub> *o mașină* [<sub>CP</sub> *care este a altcuiva*]], *dar nu sunt sigură dacă a lui Șerban<sub>i</sub>*; (\**Ana conduce* [<sub>NP</sub> *o mașină* [<sub>CP</sub> *care este t<sub>i</sub>*]])  
Ana drives a car that is GEN somebody-else but not I-am certainif Șerban’s Ana drives a car that is  
“Ana drives a car that is somebody else’s, but I am not sure if Șerban’s.”

- d. *Victima a plecat* [<sub>Adjunct</sub> *după unul dintre lingviști*], *dar nu*  
 Victim-the PAST.3S left after one of-the linguists but not  
*mi-e clar dacă dupăMirel<sub>i</sub>* (\**victimă a plecat t<sub>i</sub>*).  
 DAT.1S-be<sub>3S</sub> clear whether after Mirel victim-the PAST.3S left  
 “The victim left after one of the linguists, but it is not clear to me whether after Mirel.”

In contrast, remnants in Japanese sluices obey islands: Japanese examples analogous to the English and Romanian ones in (19) through (24) are degraded or unacceptable, as shown in (25):

- (25) a. ??*Taroo-wa* [*Akiko-to dareka*]-*o shootaisiita rasii kedo, watashi-wa*  
 Taroo-NOM Akiko-and someone-ACC invited seem but I-TOP  
*dare-o ka siranai.*  
 who-ACC Q know-not  
 “It seems that Taroo invited Akiko and someone, but I don’t know who.”
- b. \**Taroo-ga* [*Hanako-ga nanika-o katta toyuu uwasa*]-*o sinjiteiru*  
 Taroo-TOP Hanako-NOM something-ACC bought COMP rumor-ACC believe  
*ga, watashi-wa nani ka siranai.*  
 but I-TOP what Q know-not  
 “Taroo believes the rumor that Hanako bought something, but I don’t know what.” (Complex-NP Island: Merchant 1998)
- c. \**John-ga* [*dareka-ga kaite*]-*o sagasite iru rasii ga, boku-wa*  
 John-NOM someone-NOM painted-ACC looking for seembut I-TOP  
*dare-ga ka siranai.*  
 who-NOM Q know-not  
 “It seems that John is looking for a picture that somebody painted, but I don’t know who.” (Relative-Clause Island: Shimoyama 1995)
- d. \**Taroo-wa* [*dareka-ga gan kamoshirenai to-o kiita*] *naita*  
 Taroo-TOP someone-NOM cancer may-have that-ACC hear<sub>PAST</sub> because  
*kara ga, boku-wa dare-ga ka siranai.*  
 cry<sub>PAST</sub> but I-TOP who-NOM Q know-not  
 “Taroo cried because he heard that someone might have cancer, but I don’t know who.” (Adjunct Island)

The properties of English, Romanian, and Japanese sluices are summarized in Table 1:

	Overt C <sup>0</sup>	Multiple <i>wh</i> remnants	Aggressively non-d-linked <i>wh</i> words	Non- <i>wh</i> remnants	Island sensitivity
English	no	no	no	no	no
Romanian	yes	yes	yes	yes	no
Japanese	yes	yes	yes	yes	yes

Table 1: *Properties of sluices*

Romanian sluicing resembles Japanese sluicing in almost all respects, except for island sensitivity. Section 3 shows that this is a crucial difference.

### 3. *The syntax of Romanian sluices*

The facts in (26) have to be accounted for in any analysis of Romanian sluicing:

- (26) a. Multiple remnants  
 b. Non-*wh* remnants  
 c. Overt complementizers  
 d. Island insensitivity

In this section we examine possible analyses for (26). We review previous analyses of sluicing in English and Japanese and show that Romanian sluicing cannot be analyzed in terms of the latter (3.1). In section 3.2 we show that only an S-ellipsis analysis fits (26). The mechanisms behind the analysis are described in section 3.3. We discuss the implications for English in section 3.4.

#### 3.1 *Romanian sluicing as CP-ellipsis*

Because (26a-c) are facts that Romanian has in common with Japanese, it seems natural to try to extend the analysis of Japanese sluicing to Romanian. Shimoyama (1995), Merchant (1998, 2001), and Hiraiwa and Ishihara (2002) argue that Japanese sluicing should be analyzed as ellipsis of the CP-constituent of a cleft structure. A Japanese cleft consists of a copular verb, one or more NPs (the ‘foci’), and a relative-clause-like constituent (the ‘presupposition’) as in (27):

- (27) a. *Taroo-ga dareka-ni nanika-o ageta rasii ga, boku-wa*  
 Taroo-NOM someone-DAT something-ACC gave heard but I-TOP  
 [<sub>CP</sub> *Taroo-ga t<sub>i</sub> t<sub>j</sub> ageta no* ]-ga *dare-ni<sub>i</sub> nani-o<sub>j</sub> (da) ka*  
 Taroo-NOM gave<sub>PAST</sub> COMP-NOM who-DAT what-ACC is Q  
*siranai.*  
 know-not  
 “I heard that Taro gave someone something, but I don’t know to whom or what it was that he gave.”
- b. *John-ga dareka-o kubinisita rasii kedo, boku-wa [ John-ga t<sub>i</sub>*  
 John-NOM someone-ACC fired heard but I-TOP John-NOM  
*kuninisita no]-wa Bill-o<sub>i</sub> (da) to omou.*  
 fired COMP-TOP Bill-ACC is that think  
 “I heard that John fired someone, and I think that it was Bill that he fired.”

According to this analysis, the remnants of a Japanese sluice are the foci of the underlying cleft, and the elided CP is its presupposition, as in (28):

- (28) a. *Taroo-ga dareka-ni nanika-o ageta rasii ga boku-wa*  
 Taroo-NOM someone-DAT something-ACC gave heard but I-TOP  
 [<sub>CP</sub> ~~*Taroo-ga t<sub>i</sub> t<sub>j</sub> ageta no*~~] *dare-ni<sub>i</sub> nani-o<sub>j</sub>*; (*da*) *ka siranai*.  
 Taroo-NOM gave<sub>PAST</sub> COMP who-DAT what-ACC is Q know-not  
 ‘‘I heard that Taro gave someone something, but I don't know who what.’’
- b. *John-ga dareka-o kubinisita rasii kedo, boku-wa [John-ga t<sub>i</sub>—*  
 John-NOM someone-ACC fired heard but I-TOP John-NOM  
~~*kubinisita no*~~] *Bill-o<sub>i</sub>* (*da*) *to omou*.  
 fired COMP Bill-ACC is that think  
 ‘‘I heard that John fired someone, and I think that Bill.’’

As in English clefts, the focus of a Japanese cleft is in an island-sensitive dependency with a variable within its presupposition constituent. Also, Japanese clefts allow multiple pivots. As such, a CP-ellipsis analysis immediately explains the island sensitivity of Japanese clefts, as well as the availability of multiple remnants.

However, Romanian lacks clefts with multiple pivots, and in fact may lack clefts altogether (Dobrovie-Sorin 1990; Merchant 2001), so a cleft-reduction analysis *à la* Japanese will not account for (26a-c). If Romanian sluicing is not to be analyzed as CP-ellipsis, then it seems that it must be analyzed as S-ellipsis, like English sluicing. This would at least account for (26d), island insensitivity. However, as we have seen, Romanian differs from English in terms of (26a-c). In order to apply an S-ellipsis analysis to Romanian sluicing, we need to show that such an analysis can be extended to cover these facts, or to show that (26a-c) follow from independent properties of Romanian syntax.

### 3.2 An S-ellipsis account

Given that Romanian sluicing cannot be analyzed as CP-ellipsis, the remaining possibility is that it is S-ellipsis. The basic idea is that Romanian is like English, in that sluicing is ellipsis of the clausal or propositional subconstituent of an embedded question. We refer to this as S-ellipsis, rather than the more conventional IP-ellipsis, because there is still much debate as to how Romanian clause structure is to be analyzed. Alboiu (1999a, 1999b, 2000) assumes a simple CP-IP clause structure, and argues that *wh* words raise only to the specifier of IP, which is a nonargumental position in Romanian. Alboiu would then be forced to predict that Romanian sluicing is ellipsis of I', rather than IP, because the *wh* words would be in the specifier(s) of IP.

Others, including Dobrovie-Sorin (1994), Motapanyane (1998), Cornilescu (2000), and Hill (2002a, 2002b), argue that Romanian has a rich structure at the left periphery of the clause, assuming the structures proposed by Rizzi (1997), and that *wh* movement can target one of several projections that dominate IP. If we were to follow this line of argument, then Romanian sluicing could be analyzed as ellipsis of IP or of a functional projection that properly contains IP.

Because these separate analyses are based on different assumptions about which functional projections are present in Romanian, it is difficult to evaluate them with respect to one another without first resolving which of their respective assumptions are to be preferred. Because this unresolved issue is not central to our chapter we simply use the term ‘S-ellipsis,’ leaving the more appropriate label of the constituent in question to further research. We take it to be uncontroversial that it is a clause, containing a tense/mood operator, and having the semantic type of a proposition.

Whatever label we may ultimately substitute for the ‘S’ in S-ellipsis, the analysis accounts directly for three key properties of Romanian sluicing: the availability of multiple *wh* remnants, the availability of non-*wh* remnants, and the lack of island sensitivity.

### 3.3 *How the analysis would work*

We assume Merchant’s (2001) analysis of sluicing in English as a starting point. Merchant argues that S-ellipsis is licensed under semantic rather than syntactic identity. He accounts for the apparent island insensitivity of English sluicing by arguing that English sluices actually contain no (syntactic) islands. He divides island constraints into three classes, which he proceeds to explain away as being due to pragmatic, phonological, and semantic constraints, respectively. This allows sluicing to be uniformly explained as ellipsis of an S-node, with the remnants heading well-formed A’-chains rooted inside the ellipse. Semantic identity is enforced by the Focus Condition, which requires that the set of alternative propositions presupposed by the sluice entail its antecedent, and vice versa.

A crucial element of this analysis is the argument that a sluice and its antecedent have nearly identical LFs, differing only in the form and indexing of the variables they contain. Merchant (2001) assumes that focused constituents, like *wh* words, undergo quantifier raising, leaving traces inside the S-node in which they originate. Traces are interpreted as variables or E-type pronouns, which, despite being syntactically different, can have equivalent interpretations.



For example, the sluice in (1), repeated here as (29), would be (29b), with the trace bound by the *wh* word *cine*. The antecedent would be (29c):

- (29) a. [<sub>S</sub> *cineva*<sub>i</sub> [<sub>S</sub> *t<sub>i</sub> mi-a mîncat prăjiturile* ]], *dar nu știu* [<sub>CP</sub> *cine<sub>j</sub>*  
 someone me-AUX.3S eat cookies-the but not know<sub>1S</sub> who  
 [<sub>S</sub> *t<sub>j</sub> mi-a mîncat prăjiturile* ]].  
 me-AUX.3S eat cookies-the  
 “Someone ate my cookies but I don’t know who.”  
 b. [<sub>S</sub> *t<sub>j</sub> mi-a mîncat prăjiturile*]  
 c. [<sub>S</sub> *t<sub>i</sub> mi-a mîncat prăjiturile*]

The sluice in (29b) and the antecedent in (29c) differ only in the indices on the traces in their subject positions, allowing the Focus Condition to be satisfied.

3.3.1 *Multiple wh fronting*. Under the S-ellipsis analysis, the availability of multiple *wh* remnants in Romanian sluices follows directly from the fact that Romanian is a multiple *wh*-fronting language (Rudin 1988; Comorovski 1994; Dobrovie-Sorin 1990, 1994; Alboiu 2000), as shown in (30):

- (30) a. *Cine pe cine a văzut?*  
 who ACC who PAST.3S saw  
 “Who saw whom?”  
 b. \**Cine a văzut pe cine?*  
 who PAST.3S saw ACC who  
 “Who saw whom?”

S-ellipsis predicts this without further elaboration, because it would involve ellipsis of the constituent(s) below the position occupied by the fronted *wh* words. For example, (30a) can be analyzed in terms of S-ellipsis, assuming an LF representation as in (31):

- (31) *Ion* [<sub>S</sub> *cuiva<sub>i</sub> ceva<sub>j</sub> [* *a dat t<sub>i</sub> t<sub>j</sub> ]* ], *și vreau să știu*  
 Ion someone<sub>DAT</sub> something PAST.3S given and want<sub>1S</sub> SUBJ know<sub>1S</sub>  
 [<sub>CP</sub> *cui<sub>i</sub> ce<sub>j</sub> [* *a ~~dat~~ t<sub>i</sub> t<sub>j</sub> ]* ].  
 whom<sub>DAT</sub> what PAST.3S given  
 “John gave something to someone, and I want to know what to whom.”

On the other hand, if Romanian lacks clefts with multiple pivots, then one might suggest that Romanian multiple-sluicing constructions are a kind of gapping construction. Like English, Romanian has gapping and, as in English, it occurs in noninterrogative clauses and involves multiple, non-*wh* remnants:

- (32) *Gabriela a comandat o margarita și Mihai un Ursus.*  
 Gabriela PAST.3S ordered a margarita and Mihai a Ursus  
 “Gabriela ordered a margarita and Mihai an Ursus.”
- (33) *L-am văzut pe vărul meu la bibliotecă și pe nevasta sa la magazin.*  
 CL.3.M.S-PAST.1S saw ACC cousin-the my at library and ACC wife-the  
 his at store  
 “I saw my cousin at the library and his wife at the store.”

While there is still no consensus as to how gapping is to be analyzed, a gapping analysis of the Romanian data would explain both the multiple remnants and the non-*wh* phrases, because gapping applies to indicative clauses and leaves multiple remnants.

However, gapping is found in more restrictive syntactic contexts than sluicing. A sluice and its antecedent are both embedded within conjoined matrix clauses, whereas gapping only occurs between local conjunctions (Johnson 1996:21; Romero 1998:18), as in (34):

- (34) a. *Andrei a luat cartea și Marga atlasul.*  
 Andrei PAST.3S took book-the and Marga atlas-the  
 “Andrei took the book and Marga the atlas.”
- b. *\*(Cred că) Andrei a luat cartea și cred că Marga atlasul.*  
 believe<sub>1S</sub> that Andrei PAST.3S took book-the and believe<sub>1S</sub> that Marga atlas-the  
 “(I believe that) Andrei took the book and I believe that Marga the atlas.”
- c. *Cred că Andrei a luat cartea și Marga atlasul.*  
 believe<sub>1S</sub>that Andrei PAST.3S took book-the and Marga atlas-the  
 “I believe that Andrei took the book and that Marga the atlas.”

In sluicing constructions, on the other hand, the sluice and its antecedent are typically embedded inside other clauses, up to arbitrary levels of embedding as in (35):

- (35) a. *Cred că cineva a furat ceva, dar n-am nici o idee cine sau ce.*  
 think<sub>1S</sub> that someone PAST.3S stole something, but not-have<sub>1S</sub> any an idea  
 who or what  
 “I think [that someone stole something], but I don't know [who or what].”

- b. *Am auzit [ că George a spus [ că cineva vrea să-l înșele pe prietenul lui]], dar nu cred [ că știe [ cine]].*  
 PAST.2S heard that George PAST.3S said that someone wants SUBJ-CL.3S  
 cheat ACC friend-the his but not think<sub>1S</sub> that know<sub>3S</sub> who  
 “I heard that George said that someone wants to cheat on his friend but I don't think that he knows who.”

Another difference between gapping and sluicing is that sluicing allows ‘backwards ellipsis,’ meaning that the sluice precedes its antecedent in linear order, as in (36):

- (36) *Nu știu cine cu cine, dar sunt sigur că toți se vor combina cu cineva.*  
 not know<sub>1S</sub> who with who but be<sub>1S</sub> sure that everyone REFL FUT.3S combined  
 with someone  
 “I don’t know who with whom, but I am sure that everyone will get hooked up with someone.”

Gapping, on the other hand, does not allow backwards ellipsis (37):

- (37) a. *\*Andrei cartea și Marga a luat atlasul.*  
 Andrei book-the and Marga PAST.3S took atlas-the  
 “Andrei the book and Marga took the atlas.”  
 b. *\*(Cred că) Andrei a luat cartea și cred că Marga atlasul.*  
 believe<sub>1S</sub> that Andrei PAST.3S took book-the and believe<sub>1S</sub> that Marga  
 atlas-the  
 “(I believe that) Andrei took the book and I believe that Marga the atlas.”

Therefore a gapping analysis will not account for multiple-remnant sluicing in Romanian.

3.3.2 *The structure of the remnant domain and the left periphery.* As we saw, Romanian sluicing allows one or more non-*wh* remnants. Under an S-ellipsis analysis, this would follow from the presence of topicalization and focus fronting in Romanian embedded questions (38a), as well as in root clauses (38b):

- (38) a. *Nu s-a stabilit [ dacă la Bălcescu toți profesorii sunt in grevă ].*  
 not IMPERS-PAST.3S established if at Balcescu all professors-the  
 are on strike.  
 “It is not known whether all the professors at Balcescu are on strike.”

- b. *Nu știam [ că pe MARIA a ales-o Ion ] ...*  
 not knew<sub>2s</sub> that ACC MARIA PAST.3S chosen-CL.M.S Ion ...  
 "I didn't know that Ion chose Maria (rather than Ileana)..."

Topics and foci in embedded clauses are subject to the same ordering restrictions as in root clauses. First, topics must precede foci. Topics include names, definite NPs, d-linked *wh* words, and strong quantifiers (39):

- (39) a. *Mă întreb Ion cui o fi dat cartea...*  
 REFL.IS wonder<sub>1s</sub> Ion who<sub>DAT</sub> might have given book-the  
 "I have no clue who Ion might have given the book to."  
 b. *\*Mă întreb cui Ion o fi dat cartea...*  
 REFL.IS I-wonder who<sub>DAT</sub> Ion might have given book-the  
 "I have no clue who Ion might have given the book to."  
 c. *Nu știu dacă primarul pe FLORIAN îl vrea.*  
 not know<sub>1s</sub> if mayor-the ACC Florian CL.M.S want<sub>3s</sub>  
 "I don't know whether the mayor wants Florian (rather than Ion)."  
 d. *\*Nu știu dacă pe FLORIAN primarul îl vrea.*  
 not know<sub>1s</sub> if ACC Florian mayor-the CL.M.S want<sub>3s</sub>  
 "I don't know whether the mayor wants Florian (rather than Ion)."

Second, in embedded clauses as in root clauses, foci must be immediately left adjacent to the tensed verb, and to the right of any topics. Foci include *wh* words, polarity-sensitive quantifiers, negative polarity items, referential NPs pronounced with contrastive focus, and aggressively non-d-linked *wh* words. These different kinds of foci are in complementary distribution with each other (40), although multiple foci from one particular class may occur together (41):

- (40) a. *\*Nu știu [ pe cine nimeni n-a vrut să vadă ].*  
 not know<sub>1s</sub> ACC who nobody not-PAST.3S wanted SUBJ see  
 "I don't know who nobody wanted to see."  
 b. *\*Știi [ cineva pe cine vroia să lovească ]?*  
 know<sub>2s</sub> someone ACC who wanted SUBJ hit<sub>3s</sub>  
 "Do you know who somebody wanted to hit?"  
 c. *\*Mă întreb [ unde Maria trebuie să stea ( și nu Ion ) ].*  
 REFL.IS ask<sub>1s</sub> where Maria must<sub>3s</sub> SUBJ stay<sub>3s</sub> and not Ion  
 "I don't know where it is that Maria has to stay (rather than Ion)."  
 d. *\*Nu mă îndoiesc [că Maria cu nimic nu te-a deranjat].*  
 not REFL.IS doubt<sub>1s</sub> that Maria with nothing not CL.2S-PAST.3S bother  
 "I don't doubt that it was Maria that didn't bother you with anything." (adapted from Alboiu 1999a:4-5)
- (41) a. *Știi [ cine ce a mâncat ]?*  
 know<sub>2s</sub> who what PAST.3S eaten  
 "Do you know who ate what?"

- b. *Mă întreb [dacă nimeni cu nimic nu te va ajuta].*  
 REFL.1S ask<sub>1S</sub> if nobody with nothing not CL.2S FUT<sub>3S</sub> help  
 ‘‘I wonder if nobody is going to help you with anything.’’
- c. *Nu mă îndoiesc [că cineva ceva va găsi de făcut].*  
 not REFL.1S doubt<sub>1S</sub> that somebody something FUT.3S find of done  
 ‘‘I don’t doubt that somebody will find something to do.’’  
 (adapted from Alboiu 1999a:5)

This shows us that the left periphery of a Romanian embedded clause parallels the left periphery in root clauses. So, as before, an S-ellipsis analysis of Romanian sluicing directly predicts the presence of (possibly multiple) non-*wh* remnants, as these positions all c-command the ellided constituent itself. For example, (14a) is analyzed as (42):

- (42) *Mi s-a spus că [s cineva<sub>i</sub> [s t<sub>i</sub> s-a întâlnit [cu cineva]<sub>j</sub> t<sub>j</sub>]],*  
 me REFL-PAST.3S tell that someone REFL-PAST.3S meet with someone  
*mă întreb [CP dacă [s Ion<sub>i</sub> [cu Maria]<sub>j</sub> [ t<sub>i</sub> s-a ————— întâlnit t<sub>j</sub> ]]]*  
 me<sub>REFL</sub> wonder if Ion with Maria REFL-PAST.3S meet  
 ‘‘I was told that someone met with someone, I wonder if Ion with Maria.’’

### 3.4 *So what about English?*

Given our analysis, the fact that English allows only single *wh*-remnants can be explained simply by the fact that English allows fronting of only one *wh* word. However, English does allow topicalization and focus fronting, if not to the same degree as Romanian. The question is, therefore, why does English sluicing not allow non-*wh* remnants? If sluicing is simply a matter of S-ellipsis, this should be possible. One answer would be to follow Merchant (2001) in assuming that Romanian S-ellipsis is subject to the Focus Condition, as English is, but to parameterize the syntactic licensing condition, which we call the ‘S-ellipsis Condition’ (SEC). For English, the SEC requires that the ellided constituent be sister to a [+Q, +WH] complementizer. Romanian would have a more relaxed version of the SEC, which would allow S-ellipsis under sisterhood with any complementizer other than a relative-clause complementizer [+WH, -Q]. We refer to the English-type SEC as ‘strong SEC’ (43) and the Romanian-type as ‘weak SEC’ (44).

- (43) English (strong SEC):

	+Q	-Q
+WH	∅	relative <i>that</i>
-WH	<i>whether, if</i>	<i>that</i>

(44) Romanian (weak SEC):

	+Q	-Q
+WH	∅	∅
-WH	<i>dacă</i>	<i>că</i>

Another answer might be that English verbs like *know* or *wonder* have more restrictive semantics than their Romanian counterparts, which disallow topicalized constituents in their complements. This might follow from the fact that although English allows topics in root clauses, it does not allow them in embedded questions (Hudson 2003).

4. *Discussion*

According to our argument, sluicing constructions vary according to the type of ellipsis involved (CP-ellipsis vs. S/IP-ellipsis) and the number of possible remnants allowed (singleton-remnant set vs. multiple-remnant set). S/IP-ellipsis can be further subcategorized into S/IP-ellipsis with weak SEC or S/IP-ellipsis with strong SEC. There are six possible combinations represented in Figure 1. The ellipsis-type categories are placed at the top of the figure while the variation of the fronted constituent parameter lies at the bottom. Category membership tests are included within square brackets.

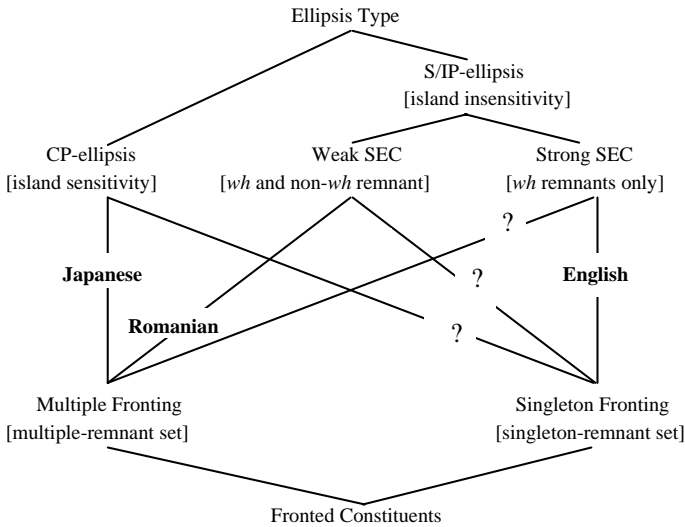


Fig. 1: *Sluicing-construction categories*

English sluicing can be described as IP-ellipsis with a strong SEC and a singleton-remnant set, Romanian sluicing as S-ellipsis with a weak SEC and a multiple-remnant set, and Japanese sluicing as CP-ellipsis with a multiple-remnant set. Future research must determine whether any languages have multiple remnants with a strong SEC, singleton remnants with CP-ellipsis, or singleton remnants with a weak SEC.

Secondly, the term ‘sluicing’ does not actually describe a syntactic configuration, but rather a correlation between a class of string languages (the surface string pronounced in a given language) and a class of semantic or pragmatic interpretations. Note that English, Romanian, and Japanese sluices are comparable in terms of their string languages in at least some cases (those involving singleton *wh* remnants) and, as far as we can tell, sluices in all three languages have comparable semantics. However, the three kinds of sluicing differ syntactically, with Romanian and English sluicing having one general kind of structure (S-ellipsis) that is disjoint with the structure of Japanese sluicing (CP-ellipsis). A subject for future research is how to relate the differences in structural description that we have seen to the similarities in string language and interpretation.

#### REFERENCES

- Alboiu, Gabriela. 1999a. “Focus Dependencies: Evidence from Romanian”. Ms., University of Manitoba.
- , 1999b. “Romanian Wh-Phrases Revisited”. Ms., University of Manitoba.
- , 2000. *The Features of Movement in Romanian*. Ph.D. dissertation, University of Manitoba.
- Comorovski, Illeana. 1994. *Interrogative Phrases and the Syntax-Semantics Interface*. Boston: Kluwer.
- Cornilescu, Alexandra. 2000. “The Double Subject Construction in Romanian”. *Comparative Studies in Romanian Syntax* ed. by Virginia Motapanyane, 83-133. Dordrecht: Elsevier.
- Dobrobie-Sorin, Carmen. 1990. “Clitic Doubling, Wh-Movement, and Quantification in Romanian”. *Linguistic Inquiry* 21:3.351-397.
- , 1994. *The Syntax of Romanian*. Berlin: de Gruyter.
- Hill, Virginia. 2002a. “Adhering Focus”. *Linguistic Inquiry* 33:1.164-172.
- , 2002b. “Complementizer Phrases in Romanian”. *Rivista di Linguistica* 14:2.
- Hiraiwa, Ken & Shinichiro Ishihara. 2002. “Missing Links: Cleft, sluicing, and the no-da construction in Japanese”. *MIT Working Papers in Linguistics* 43.35-54.
- Hudson, Dick. 2003. “Trouble on the Left Periphery”. *Lingua* 113.607-642.
- Johnson, Kyle. 1996. “In Search of the English Middle Field”. Ms., University of Massachusetts at Amherst.
- Lobeck, Anne. 1995. *Ellipsis: Functional Heads, Licensing, and Identification*. Oxford: Oxford University Press.

- Merchant, Jason. 1998. "Pseudosluicing: Elliptical clefts in Japanese and English". *ZAS Working Papers in Linguistics 10* ed. by Artemis Alexiadou, Nanna Fuhrhop, Paul Law, & Ursula Kleinhenz, 88-112. Berlin: Zentrum für Allgemeine Sprachwissenschaft.
- , 2001. *The Syntax of Silence: Sluicing, islands, and the theory of ellipsis*. Oxford: Oxford University Press.
- Motapanyane, Virginia. 1998. "Focus, Checking Theory, and Fronting Strategies in Romanian". *Studia Linguistica* 52:3.227-243.
- Pesetsky, David. 1987. "Wh-in-Situ: Movement and unselective binding". *The Representation of (In)Definiteness* ed. by Eric Reuland & Alice ter Meulen, 98-129. Cambridge, MA: MIT Press.
- Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery". *A Handbook of Generative Grammar* ed. by Lilliane Haegeman, 281-337. Dordrecht: Kluwer.
- , 1998. *Focus and Reconstruction Effects in Wh-Phrases*. Ph.D. dissertation, University of Massachusetts at Amherst.
- Ross, John. 1967. *Constraints on Variables in Syntax*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Rudin, Catherine. 1988. "On Multiple Questions and Multiple Wh-Fronting". *Natural Language and Linguistic Theory* 6:4.445-501.
- Shimoyama, Junko. 1995. "On 'Sluicing' in Japanese". Ms., University of Massachusetts at Amherst.





**ROMANCE INTONATION FROM A COMPARATIVE  
AND DIACHRONIC PERSPECTIVE  
POSSIBILITIES AND LIMITATIONS**

JOSÉ IGNACIO HUALDE

*University of Illinois at Urbana-Champaign*

**0. Introduction**

Quite clearly, the considerable depth of accumulated knowledge on diachronic Romance phonology does not extend to the realm of intonation. Those familiar with the available textbooks on comparative Romance linguistics and manuals devoted to the history of individual Romance languages will be aware that, whereas other aspects of historical phonology may be treated in great detail in such books, one does not expect to find a chapter or even a short section on the diachronic evolution of the intonational system from Latin to Romance.

At most we may find some vague statements to the effect that Classical Latin had some ill-understood musical accent that, for ill-understood reasons, was lost in all Romance languages. Statements like this can be found in many books, both old and recent. Pei (1976:64), for instance, tells us the following: “The real problem of the Classical Latin accent lies in its essential nature. It shares with its older sisters, Greek and Sanskrit, the feature of being based on musical pitch rather than the stress of the voice that is natural with us of the modern Western world.”

Even more recently, we find restatements of this traditional view. For example, in his otherwise excellent history of the Spanish language, Penny (1991) puts forth this opinion:

It is thought that early Latin had a type of accent in which pitch was the predominant element.... However, for reasons that remain obscure, spoken Latin underwent a change of accent-type and came to have an accent in which energy-deployment dominated. This type of accent (stress-accent) is the one which continues to characterize the majority of Romance languages (including Spanish) and is also the type used by English. (Penny 1991:35)

As I argued in Hualde (2003), there is very little evidence for this position (see also Baldi 1999). It is far more likely that such a transformation never took

place.<sup>1</sup> As Posner (1996:99) points out, “some scholars believe that the tonal accent, described by Latin grammarians, may in fact have been an artifact of the literary language, imitated from prestigious Greek.” But Classical Greek itself may have differed only minimally from Modern Greek (and the Romance languages) in the realization of accent. Classical Greek, like modern Serbo-Croatian, Swedish/Norwegian, and some Low German dialects, most likely was a stress-accent language, which, in addition, had a lexical contrast in the alignment of accentual peaks. In Greek this contrast was limited to syllables with a long vowel or a diphthong, as in some of the other European languages that have a contrast of this type (early peak, on the first mora, vs. late peak, on the second mora). Modern Greek has lost this lexical contrast, but stressed syllables still have the important function of serving as anchoring points for the alignment of intonational pitch accents. This is also true of the Romance languages and, we may assume, Latin. The traditional distinction between ‘expiratory’ and ‘musical’ accent does not seem to have much factual basis.

The reasons why such poorly grounded opinions persist and virtually no headway has been made in the study of the origin and diversification of the Romance intonational systems have to do with difficulties inherent to the study of intonation. First of all, the existence of a rich philological record, which has given the field of Romance historical linguistics such an advantage over the study of other language families, is of no help in this particular case. We cannot tell much about intonation from the available written sources. Second, the other traditional strength of diachronic Romance linguistics is the availability of many detailed descriptions of contemporary varieties, permitting a solid application of the comparative method. But, here again, the available descriptive work on the intonation of Romance languages and dialects is less abundant and less reliable than that concerning other phonological aspects. For these two reasons, I believe, the field of diachronic Romance ‘intonology’ (my coinage) has remained underdeveloped until now.

Whereas the nature of the philological evidence on intonation has not changed and obviously is not likely to change, the last few decades have witnessed considerable progress in the study of intonation in living languages, due in no small measure to technological advances. Currently, much effort is being devoted to the study of the intonational systems of Romance languages and dialects and this is bearing fruit in the form of increasingly more accurate and complete descriptions. Using the comparative method to reconstruct aspects of the evolution of Romance intonation seems a much more feasible

---

<sup>1</sup> Many of the topics discussed in this chapter are also touched on in Hualde (2003), although an attempt has been made to avoid repetition.

task now than it was before and perhaps efforts in this direction would not be completely wasted.

We must, nevertheless, be aware of certain special problems for establishing genetic relationships in matters intonational:

(i) The comparative method relies on the identification of cognates: items with comparable form and meaning, such as It. *cane*, Port. *cão*, and Fr. *chien*, all meaning “dog.” One problem in our case is that identifying what count as comparable intonational elements is particularly difficult. This applies to both aspects of the equation: form (the shape of contours) and meaning (the meanings that are conveyed by them). Intonation involves the pragmatic use of pitch, and pragmatic meaning is a notoriously thorny issue. In this respect, we may note that in some other European languages where pitch has in part a lexically contrastive function, such as Swedish/Norwegian and Rhenanian Low German/Dutch dialects, historical reconstruction operates on much firmer ground and some interesting diachronic proposals on the interaction between accent and intonation have indeed been put forward (see Riad 1998, 2000, 2003 for Scandinavian; Gussenhoven 2000 for Rhenanian Germanic).

(ii) For the comparative method to work, the link between the two parts of the sign, form and meaning, must be arbitrary. Regarding intonation, however, there are clear universal tendencies that make the application of the technique less trustworthy. As Ladd (1996:113) notes, common intonational phenomena of cross-linguistic validity include (a) the use of a falling contour to indicate finality and a high or rising tone to indicate nonfinality (i.e., either incomplete statements or questions, where an answer is expected), (b) the progressive reduction and declination of pitch excursions from the beginning to the end of the utterance in declaratives, and (c) the marking of new or important information by means of localized pitch movements. But, in spite of these cross-linguistic tendencies, Ladd makes a compelling case for the language specificity of intonational structure. Some aspects of the intonational system are unquestionably language specific. Even the cross-linguistic phenomena just noted are all tendencies rather than hard universals and exceptions to all of them can be found. Declaratives with a final rise are found both in some English varieties and in Dominican Spanish (Willis 2003), and falling contours in yes/no questions are typical of Caribbean Spanish (Quilis 1987; Sosa 1999) and of some Italian dialects (Grice 1995; D’Imperio 2002).<sup>2</sup> In some Basque dialects the presence versus absence of pitch accents is a lexical property of words and cannot be manipulated for pragmatic purposes (Hualde et al. 2002;

---

<sup>2</sup> Grice (1995) argues that the Palermo Italian interrogative contour does not violate the universal high or rising tone in questions because there is a rising accent before the final fall.

Elordieta & Hualde 2003). Nevertheless, the existence of the above-mentioned tendencies is undeniable. In practical terms, what this means is that it will be easier to argue for common inheritance when we are dealing with intonational phenomena that are clearly marked or even go against universal tendencies than when we observe a coincidence in facts that are cross-linguistically so widespread that common origin does not constitute the best hypothesis to explain the similarity.

(iii) An added difficulty is distinguishing inherited from borrowed elements. Intonational contours appear to be particularly susceptible to borrowing. In the past, there have been claims of wholesale borrowing of intonational structure. For instance, Henríquez Ureña (1938:335) claimed that, at the lower socio-economic levels, Mexican Spanish intonation is identical to Nahuatl intonation (cf. O'Rourke, this volume). Recent work has demonstrated that Romanian shares some intonational features with other Eastern European languages (Ladd 1996; Grice, Ladd, & Arvaniti 2000). Kaisse (2001) argues that a specific contour in Argentinean Spanish has been borrowed from Italian. O'Rourke shows that Cusco Spanish intonation has likely been influenced by the Quechua intonational system in certain specific features. Elordieta (2003) explores the possible influence of Basque prosody in the Spanish speech of bilingual speakers from Lekeitio.

Briefly, then, the comparative method is applicable to the study of intonation to the extent that we can identify 'cognates,' that is, intonational elements with comparable forms, comparable meanings, and idiosyncratic form-meaning mapping, where the best explanation for the correspondence across the languages is inheritance from a common origin, rather than intonational universals or borrowing. Being aware of possible pitfalls, I believe we can approach the comparative and diachronic study of Romance intonation with guarded optimism.

This chapter is organized as follows. In section 1, I discuss some common Romance patterns that have been pointed out in the literature or that emerge from the comparison of studies on various Romance languages. In section 2, I extend the comparison to Occitan, a Romance language that has been neglected in intonational studies but that seems to me very important for the comparative and diachronic study of Romance intonation because, geographically, Occitan forms a link between French, the most innovative Romance language in prosodic matters, and both Ibero-Romance and Italo-Romance (on Occitan prosody cf. Meisenburg 2001).

## 1. *Common Romance prosodic features*

### 1.1 *Nuclear-accent assignment rules*

At this stage in the investigation, we may ask whether any prosodic features can be identified that appear to fulfill the two requirements of being common to a number of Romance languages and being peculiar to Romance, especially when compared to its closest neighbors. Perhaps one such feature is the patterns of nuclear-accent assignment in certain constructions. Several authors have remarked on a difference in this respect between English and other Germanic languages, on the one hand, and a number of Romance languages on the other. Ladd (1996:191-192) points out that, whereas in examples such as (1a), with a NP + infinitive sequence, nuclear stress would be placed on the noun in English (or German) in the unmarked case, in Italian (or Spanish) the corresponding unmarked contour would display main prominence on the infinitive, as in (1b). Prieto (2002b:71-72), agreeing with Ladd, shows that Catalan behaves like Italian in this respect (1c). The same pattern is also found in Spanish (as noted by Ortiz-Lira 1994 and Ladd), for which I provide my own example in (1d):

- (1) a. English: They gave him a TUNE to play.  
 b. Italian: *Gli hanno dato una musica da SUONARE.*  
 c. Catalan: *Li van donar una tonada per TOCAR.*  
 d. Spanish: *Le han dado una melodía para TOCAR.*

To the extent that it is a feature that distinguishes Romance from other neighboring non-Romance European languages, such as the West Germanic languages, this property can be reconstructed for an earlier, common Romance stage. Nevertheless, we may note that it is the West Germanic languages that appear to be special in this respect, in that they differentiate between nouns and verbs in their relative accentability. The Romance languages simply lack this contrast and have a general rule assigning nuclear accent to the last content word in broad-focus utterances, which applies in noun + infinitive constructions as well.

Ladd (1996:175-184) discusses several other differences in patterns of prominence between West Germanic (English, Dutch, German) and Italian (and Romanian). In general, what we observe is a stronger tendency in the Romance languages to place the nuclear accent on the last content word (word order being exploited for pragmatic purposes to a much larger extent than in Germanic). Ortiz-Lira (1994), in his detailed contrastive analysis of accent placement in Spanish and English, also points out this tendency to avoid deaccentuation of constituents in final position in Spanish, even in the case of

given or repeated information, noticing that, in this respect, Spanish behaves much more like Italian, Romanian, and Portuguese. Catalan also shows nuclear-accent patterns similar to those of the other Romance languages, and different from the English ones (Vallduví 1991, 1992).

Prosodic features found in a number of Romance languages but not in languages belonging to other families, especially neighboring families such as Germanic, are potential candidates for extrapolation to an earlier, common stage, as mentioned before. Features shared by languages that are not in immediate geographical contact are particularly interesting in this respect. As noted earlier, Ladd (1996) has shown that Romanian presents some intonational features that are found in languages of Eastern Europe but not in other Romance languages. Romanian is in fact an important language for reconstructive purposes because of its geographical isolation from the rest of Romance. According to Ladd (1990, 1996) Romanian is like Italian and unlike West Germanic languages in its rules for the placement of nuclear stress in declaratives, whereas it resembles other Eastern European languages such as Russian, Hungarian, and Greek in placing the main accent on the verb in yes/no questions, a context where the general rule is the same in West Germanic and the other Romance languages. If further work on Romanian were to identify other features that Romanian shares with other Romance languages but that, on the other hand, are not found in Hungarian or Slavic, there would be a strong case for accepting that those shared features are the product of inheritance.

### 1.2 *Shape of nuclear accents and broad- versus narrow-focus contrast*

Although the connection is not explicitly made by Ladd (1996), a fact that favors the consistent placement of nuclear accent on the last content word in Italian is that this language has a contrast between two possible nuclear accents, with a difference in pragmatic meaning (D'Imperio 2002): Whereas broad-focus sentences have a falling nuclear accent (H+L\* = a fall on the stressed syllable from a peak on the pretonic), narrow focus on the last word is conveyed by a different pitch accent, a rising accent with a tonal peak on the stressed syllable (analyzed by D'Imperio as L+H\*). As D'Imperio notes, the existence of this choice of pitch accent allows for a contrast in interpretation that is not predicted by theories based on English, which, instead, predict neutralization between broad-focus sentences and sentences with narrow focus on the last lexical item.

This contrast between two nuclear accents, associated with broad and narrow focus, is also found in other Romance languages. Figure 1 compares

schematics for broad- and narrow-focus statements in Italian (redrawn from D'Imperio 2002) and in European Portuguese, with the one-word minimal pair that Frota (2002) gives as an example. A very similar contrast has been described for Catalan (Prieto 2002a, 2002b).

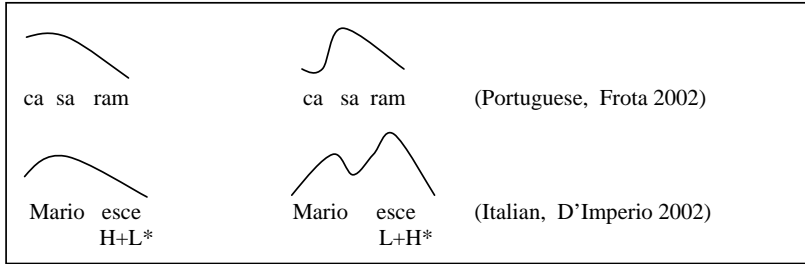


Fig. 1: Nuclear broad- and narrow-focus contours in European Portuguese and Italian

Based on available descriptions for a number of Romance languages, we can reconstruct a common Romance intonational system with three distinct pitch accents in declarative utterances: a prenuclear accent, characterized by a rise through the stressed syllable with a peak generally displaced to the post-tonic,<sup>3</sup> and two nuclear accents, a broad-focus nuclear accent, which is a fall through the stressed syllable, and a narrow-focus nuclear accent, with a rise that reaches a maximum within the stressed syllable before falling. This is illustrated in Figure 2. (To refer to these three accentual shapes I will use the terms 'rising,' 'falling,' and 'circumflex,' respectively.)

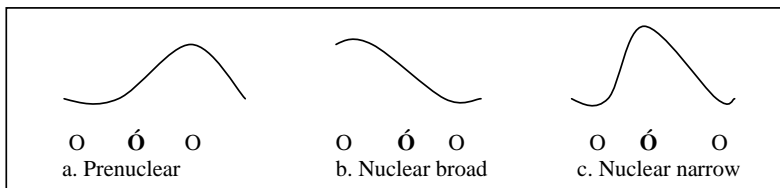


Fig. 2: Common Romance inventory of pitch accents in declarative sentences

A stricter alignment of accentual peaks (or complete rise-fall movements) with stressed syllables to convey narrow-focus prominence is a cross-linguistic

<sup>3</sup> Estebas-Vilaplana (2003) demonstrates that, at least in Central Catalan, the exact position of the peak is a function of the distance between the stressed syllable and the end of the word. Consistent with this claim, the greatest displacement of the peak with respect to the stress syllable takes place in proparoxytones and the least displacement in oxytones.



tendency (see, e.g., Smiljanic 2002 for Serbo-Croatian). This may result in the retraction of the peak in the case of non-phrase-final words, where otherwise the peak would be displaced to the post-tonic. In the case of phrase-final words, what we find in Romance is the choice of a different contour, with apparent 'peak protraction' in phrase-final words. We can thus see this choice of accentual contours as a language-specific instantiation of a cross-linguistic intonational tendency.

As mentioned, broad-focus nuclear contours involving a fall throughout the stressed syllable, as in Figure 2 (b), are quite consistently found in Italian, European Portuguese, and Catalan. Such nuclear contours are also common in some Spanish dialects, most notably Chilean (cf. Ortiz-Lira 1994; Beckman et al. 2002). They are, however, less frequent in some other Spanish varieties. Face (2002) does not describe this contour in his study of focus marking in Madrid Spanish. Neither is this falling nuclear pitch accent found in the data produced by the bilingual Basque-Spanish speakers in Elordieta (2003). It is possible that the absence of the falling nuclear contour in the data reported in these studies has something to do with the fact that what is being investigated is read speech, as opposed to spontaneous conversation. Perhaps a clearer exception is found in Mexican Spanish, where declarative sentences very frequently display a 'circumflex intonation' in contexts that are not particularly emphatic and without implying narrow focus on the final constituent (Quilis 1987:134). This circumflex intonation is a nuclear configuration with a peak on the stressed syllable as in Figure 2 (c), which is upstepped with respect to preceding accentual peaks. The broad shape of this circumflex contour is that used for narrow focus on the last word in Italian, for example, but the Mexican contour is used outside of narrow-focus contexts. Thus, it is clearly not the case that all Romance varieties systematically observe the contrast illustrated in Figure 2. Given the geographical extension of the contrast between a broad-focus nuclear falling contour and a circumflex narrow-focus contour with peak in the stressed syllable, we are necessarily led to the conclusion that those Spanish (and possibly other) dialects that lack this pragmatically determined pitch-accent contrast must have lost it or reinterpreted it, rather than the other Romance languages having innovated.

On the other hand, a contrast similar to that reported for Italian, Catalan, Portuguese, and so on may be found in some German dialects. In particular, a broad-focus declarative contour consisting of prenuclear L+H\* accents followed by a nuclear falling H\*+L accent is described for Southern German by Truckenbrodt (2002). As noted earlier, to the extent that common Romance patterns are also found outside of Romance, common inheritance loses force as

the best account for the similarity (although in this particular case, depending on the geographical extension of the pattern within Germanic, borrowing from Romance may still be a reasonable hypothesis). An important difference is that final verbs in Southern German are unaccented (consistent with the description in Ladd 1996 discussed before), and the falling nuclear accent is associated with the preceding noun in verb-final sentences, whereas in Romance the last content word in the sentence receives the nuclear accent. The overall contour is thus different in Romance and Southern German, even if the shape of pitch accents is similar, because there is a difference in the pattern of accent placement.

Finally, variation in prenuclear accents is also found across Romance. For instance, Caribbean dialects of Spanish often present a very late onset of the accentual rise. Whereas in most Romance varieties (and Greek, cf. Arvaniti, Ladd, & Mennen 2000) the rise in prenuclear accents typically starts immediately after a valley at the beginning of the stressed syllable (the peak being reached on the post-tonic), in Caribbean Spanish there is often a flat contour throughout most of the stressed syllable and the rise only starts toward the end of this syllable (Beckman et al. 2002; Willis 2003). These Caribbean prenuclear accents are thus similar to the Serbo-Croatian rising accent (as described in Smiljanic 2002, among others).

One of the most striking deviations with respect to the common Romance pattern of prenuclear accents is found in European Portuguese. Frota (2002) points out that, in this language, most typically in multiword declaratives, there is only one rise corresponding to the first stressed syllable in the sentence and a flat contour up to the nuclear fall. Stressed syllables in medial position do not receive tonal relief. As Frota remarks, European Portuguese differs from Spanish, Catalan, and Italian in its very sparse use of prenuclear accents. Since Brazilian Portuguese sides with the other Romance languages in this regard, this is clearly an innovation in European Portuguese. It is likely that this intonational innovation is linked to the process of reduction and deletion of unstressed vowels in European Portuguese, an aspect in which this language also differs from both Spanish and Brazilian Portuguese.

## 2. *From common Romance to Gallo-Romance*

Impressionistically, at least, the most divergent Romance language in intonational respects is clearly French. When we get to the specifics of the analysis, how different French intonation really is from other Romance intonational systems is, to some extent, a matter of opinion. Jun and Fougeron (1995, 2000, 2002) analyze French intonation as fundamentally different in its

prosodic structure from standard analyses proposed for other Romance languages within the same basic analytical framework. Post's (1999, 2000, 2002) proposal, on the other hand, makes French prosody more compatible with other analyses of Romance intonation. But, in any case, the fact remains that French is clearly distinct from the rest of the major Romance languages.

As I argued in Hualde (2003), in order to investigate how the French intonational system has evolved and diverged from the common Romance one, it makes good sense to pay some attention to Occitan, which forms a bridge with both Italo-Romance and Ibero-Romance and is likely to represent an intermediate stage concerning intonational features.

### 2.1 *Initial rises and final falls in Occitan declaratives*

A preliminary analysis of intonational contours in narratives shows that Occitan shares basic prosodic properties with its southern neighbors. As in Catalan and Italian, declarative utterances may contain one or more rising prenuclear accents and typically a falling nuclear accent. One feature in which Occitan differs from Ibero- and Italo-Romance, on the other hand, is in the frequency of tonal rises not aligned with a stressed syllable but, rather, with the beginning of prosodic words, even on syllables immediately preceding the lexically stressed syllable of the word. My impression is that these initial rises are prominence-lending tonal events, and thus classifiable as pitch accents, but it is possible to remain neutral in the description (i.e., whether these rises are better analyzed as pitch accents or as boundary tones). Some illustrative examples of this phenomenon are shown in Figures 3 through 5. The data have been obtained from an analysis of the recordings that accompany Loddo (1993), a collection of folk stories in a Languedocien variety. This is the same source employed in Hualde (2003), though different examples are analyzed here. The pattern is relatively frequent and all three examples given in Figures 3 through 5 were taken from a short fragment of the same folk narrative transcribed by Loddo (1993:82).

I indicate lexically unstressed syllables bearing pitch accents (secondary accent) in bold in orthographic transcription. Syllables possessing lexical stress are underlined. In phonetic transcription, initial, nonlexical, or secondary accents are indicated with (") before the syllable carrying the pitch accent, and syllables with lexical stress that actually receive prominence are indicated with the IPA primary stress symbol (').

In Figure 3 (*Deu pas èsser menut tampauc!* [dew par εse "me'nyt ta'pawk] "I bet he is no weakling either!") there is a clear pitch excursion over the first syllable of *menut*, even though lexical stress is on the next syllable.

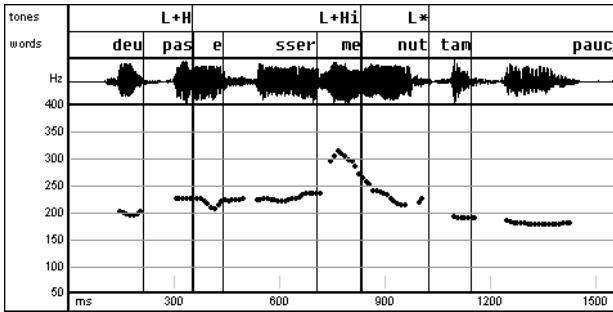


Fig. 3: Word-initial prominence on lexically unstressed syllable (menut)

In Figure 4 (*per far de redondials* [per 'fa ðe 'reðun'djals] “to make harness rings”) we can see the same type of tonal event on the first syllable of *redondials*.

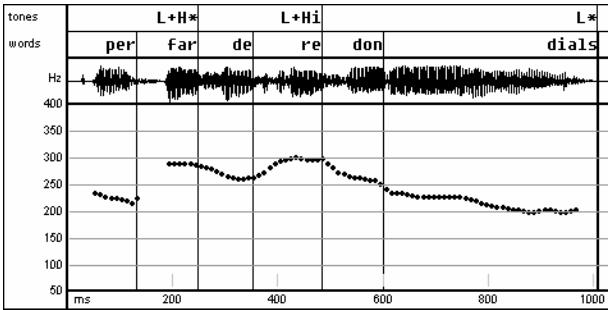


Fig. 4: Word-initial prominence (redondials)

In Figure 5 (*Ne trapan un que fasiá a las bòlas ambe des mòlas de molin* [ne 'trapo 'yn ke fazjo a laj 'βòloz ambe ðe 'mòloj 'ðe mu'li] “They find someone who was playing marbles with mill-stones”), there is also tonal prominence on the first syllable of the last prosodic word, *de molin*, the preposition *de*, from which there is a drop to the end of the sentence. This last example illustrates a related property of Occitan: Final nuclear falls, instead of starting from the pretonic as in other Romance languages, may have their starting point further to the left, on a peak on the initial syllable of the prosodic-word unit (including clitics, as in this example).

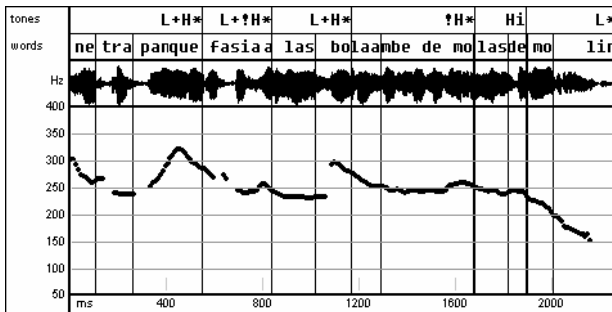


Fig. 5: Final fall from syllable with secondary accent

As in the other Romance languages, continuation in Occitan is indicated by an incomplete fall. Figures 6 and 7 constitute a single textual unit, where the narrator assumes the role of one of the characters in the story giving instructions to the other characters: “*Es pas complicat: la vos cal tornar tampar e se manifestarà quand aurà pus deguns de sa familha sus terra*” (Loddo 1993:262) [es pas kumbli'kat M% "laβus'kal tur'na tam'pa M% e "se manifesta'ra M% kand aw'ra pas 'py 'ðy ðe sa fa'miʎo syr 'tero L%] “It is not complicated: You must entomb her again, and she will manifest herself when there is no longer anyone from her family on Earth.” Each of the three incomplete declaratives has a continuation rise and the whole sequence ends with a fall throughout the last stressed syllable.

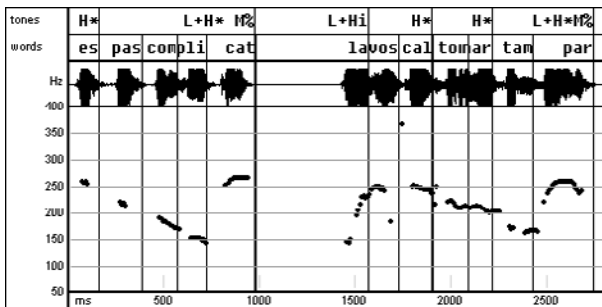


Fig. 6: Two incomplete declarative sentences with continuation rise

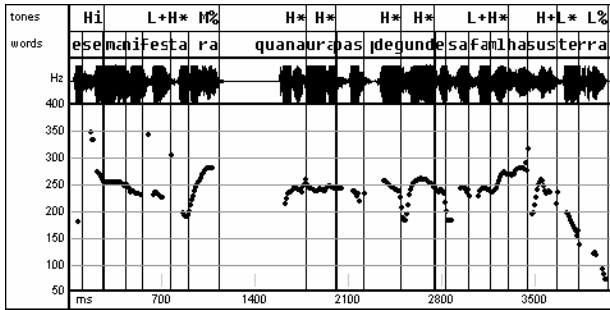


Fig. 7: Second part of narrative, continuation rise (L+H\* M%) and final fall (H+\* L%)

The frequent presence of localized pitch excursions on syllables without lexical prominence (secondary or nonlexical accents) appears to be a feature in which Occitan differs even from languages as closely related to it as Catalan. On the other hand, this is a clear feature that it shares with French. This is in fact an important intonational respect in which French—together with Occitan, we may now add—differs from the other Romance languages. Post (2000:82-83), agreeing with previous descriptions of French prosody, states that in French “there is a strong tendency for phrases to be marked by initial pitch prominences, resulting in a hammock shape of the pitch contour. That is, when there is enough room for it to be realized, a secondary accent will surface as early as possible in the phrase.” As Oakes (2002) remarks, in certain styles it is also very frequent to have secondary stresses even immediately before the syllable with primary or lexical accent, as in his example *l’envoie* [l’ã<sup>1</sup>vwa], produced with two pitch excursions in his Figure 4 and perfectly comparable to the example *plaser* [plã<sup>1</sup>zɛ] in one of the Occitan illustrations in Hualde (2003:192). Languages like Spanish, Catalan, and Italian can also make use of initial secondary stresses to emphasize certain words in the utterance, but they do so with much less frequency and not on immediately pretonic syllables (cf. Ortiz-Lira 2000 for Spanish).

The fact that Occitan shares this feature with French is interesting. French, as is well known, historically lost the lexically contrastive use of stress that is displayed by the other Romance languages (as in, e.g., Sp. *término* “conclusion, boundary,” *termino* “I finish,” *terminó* “s/he finished”; It. *càpito* “I turn up,” *capító* “understood,” *capitò* “s/he turned up”). One could surmise that the lack of lexical contrast in the location of stressed syllables is what allows French more freedom in the placement of local tonal events. One could argue that tonal excursions are not necessarily aligned with stressed syllables in French, precisely because stress is not lexically contrastive. Loss of contrastive

stress may have thus triggered a major change in the intonational system, in this view. The problem for this theory, we see now, is that Occitan shares the use of initial accents with French, even though Occitan does have lexically contrastive stress (final vs. penultimate), since vowels in final syllables have not been weakened in Occitan. Lack of contrastive stress is thus not a prerequisite for the proliferation of word-initial tonal rises.

Something that is probably impossible to determine with certainty is whether what we see in modern Occitan represents an intermediate stage in the evolution of French. Finding this pattern in Occitan does not allow us to completely rule out the possibility that French developed its initial rises after it had lost contrastive stress and that Occitan has recently acquired this feature under the influence of French. Nevertheless, there is some independent evidence that this is indeed an old prosodic property of French, antedating the weakening of final syllables. From the examination of the distribution of word boundaries within verse lines, Penson (1993) concludes that, already in Racine and even the *Chanson de Roland*, there was typically “accent on the counter-tonic of oxytones of more than two syllables and paroxytones of more than three” (Penson 1993:24).

## 2.2 Interrogative patterns

In Romance languages, as in English, neutral pronominal questions generally have a final falling contour, from a high point on the question word. This is also true of Occitan. An example is given in Figure 8 (*E ont anatz como aquò?* [e 'unt a'nas kumo'kɔ] “Where are you<sub>PL</sub> going like that?”).

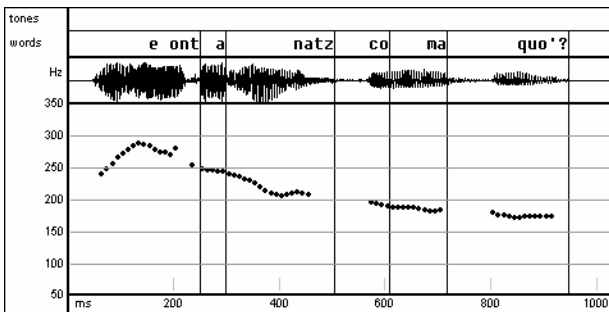


Fig. 8: Pronominal question with final fall

In most Romance languages, yes/no questions do not commonly display any special morphological or syntactic marking and are distinguished from statements solely by intonational means. Thus, for instance, Grice and Savino (2003) point out that an Italian text such as *vado a destra* can be a statement “I go to the right,” a ‘query’ or request for information “Do I go to the right?,” or a request for confirmation or ‘check’ “(So) I go to the right?.” The distinction among these pragmatic meanings can only be intonationally expressed. The same is true of a Spanish example such as *Llegaron tus amigos* “Your friends arrived,” “Did your friends arrive?,” “(So), your friends arrived?” and examples could also be given for the other Romance languages.

Queries normally have ‘circumflex’ patterns (for Spanish cf. Quilis 1993). This configuration, a rise and fall with peak aligned with the stressed syllable, has basically the same shape as the narrow-focus nuclear-accent contour. Figure 9 provides an interesting illustration. The sentence *Irosa seràs dins ton maridatge* [i'ruzo se'raj ðin 'tun ma'rjatse] “You will be happy in your marriage” has emphatic narrow focus over the first word and we can see a clear peak on its stressed syllable, followed by a fall. The utterance *Irosa?* [i'ruzo] “Happy?,” which has the status of a ‘check’ in the fictional dialogue within the narrative, can be seen to have the same contour, but with a greatly expanded range.

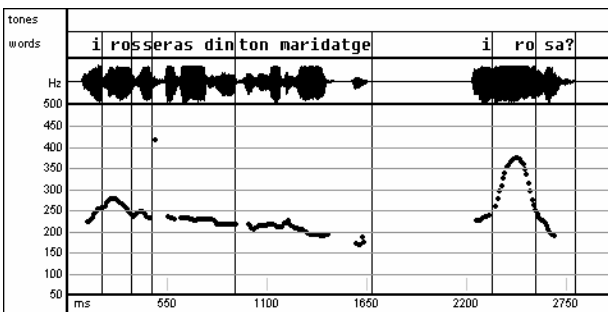


Fig. 9: Statement and following check yes/no question

All the Occitan tokens of yes/no questions that I have been able to inspect all have a circumflex ending. An example is Figure 10 *Venes pas ne tastar un bocin?* [ˈbenes pa ne tasˈta (y)mbuˈsi] “Aren’t you coming to taste a bit of it?”



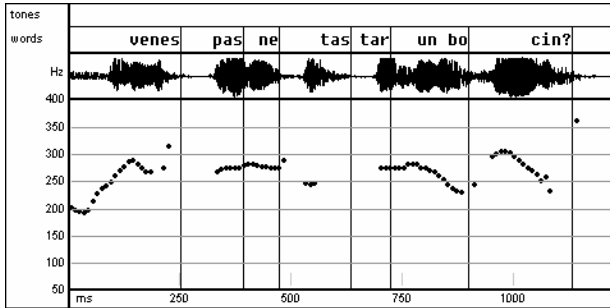


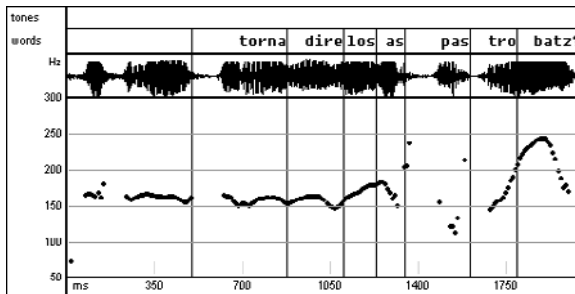
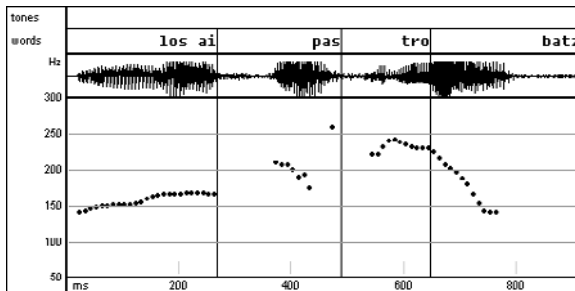
Fig. 10: *Yes/no question with circumflex contour*

In most Romance varieties, yes/no questions are differentiated from statements by a final rising contour, allowing for a clear distinction between question and statement in the absence of morphological or syntactic cues (e.g., Quilis 1993:429; Sosa 1999:153 for Spanish; Avesani 1995 for Italian; Di Cristo 1998:202 for French; Frota 2002 for European Portuguese). Catalan, in addition to distinguishing questions from statements only by their rising terminal contour, as in *Veureu la Maria?* ‘‘Will you see Maria?,’’ can mark the interrogative function with the word *que* in sentence-initial position, as in *Que veureu la Maria?*, in which case the final contour may be of the falling type (Prieto 2002a:422). Interestingly, though, morphosyntactically unmarked questions with a final fall (a circumflex contour), very much like those of Occitan, have been described for a number of Spanish and Italian varieties as well as for Brazilian Portuguese.

Within Spanish, circumflex contours in questions are typical of Caribbean dialects (Quilis 1987; Sosa 1999; Beckman et al. 2002). Both Quilis and Sosa, in fact, suggest that a final rise is not possible in normal yes/no questions in Puerto Rico Spanish and other Caribbean dialects.<sup>4</sup> In Italy, circumflex yes/no interrogative contours have been reported for all southern regional varieties of Italian that have been described in some detail (cf. Grice 1995; D’Imperio 2002, among others), whereas more northern standard Italian is said to employ a rising contour (Avesani 1995). As for Brazilian Portuguese, de Moraes (1998) describes circumflex contours with a rise on the last stressed syllable and a final fall as the usual intonation of yes/no questions in this Portuguese variety, without specifying regions (cf. also Quilis 1988).

<sup>4</sup> Nevertheless, Willis (2003) reports frequent rising contours in questions for the Dominican dialect of Santiago de los Caballeros, in addition to the falling Caribbean contours reported by other authors.

Naturally, a question that arises is how yes/no interrogatives are distinguished from statements, particularly, from statements with narrow focus on the last word, when final contours of the circumflex type are used in both types of sentences. The contrast with broad-focus declaratives is, on the other hand, straightforward in varieties where these sentences have a final falling accent. This is illustrated with a near minimal pair from the same narrative (Loddo 1993:52). In Figure 11, the interrogative sentence (*Lo patron ie tòrna dire: los as pas trobats?*) [luz as pas tru'βats] “(The boss says to him again:) Haven’t you found them?” has a final circumflex accent over the last syllable of the sentence, which carries primary stress. In the near minimally contrastive declarative in Figure 12, *los ai pas trobats* [luz aj pas tru'βats] “I haven’t found them,” however, the final stressed syllable *-bats* bears a falling contour, quite different in the alignment of peak and fall from the final contour in Figure 11. Although both contours have a boundary fall, the type of nuclear pitch accent differentiates yes/no questions from neutral or broad-focus declaratives.

Fig. 11: *Interrogative*Fig. 12: *Declarative*

Possible ambiguity would arise between circumflex yes/no questions and declaratives with a circumflex accent, which, as discussed in section 1, is typical of declaratives with narrow focus on the final constituent. Regarding Neapolitan Italian, D'Imperio (2002) states that, although the final shape of questions is very similar to that of statements with narrow focus on the last word, a rise-fall contour in both cases, the alignment of the tonal events is different in the two cases: "Specifically, the entire rise-fall appears to be timed later (relative to the stressed vowel) in questions than in statements" (D'Imperio 2002:44). As for Puerto Rican Spanish, Quilis (1987) suggests that questions may have a wider range than statements. This is also true of the Occitan recordings that I have examined. The difference in range was illustrated in Figure 9. Quilis also claims that Puerto Rican questions sometimes show incomplete falls, but, of course, incomplete falls are also found in incomplete statements. More careful analysis may very well also reveal differences in alignment in other Romance languages with typically circumflex yes/no questions like those found by D'Imperio for Neapolitan Italian (see also Garrido 1991:64; Cid Uribe & Ortiz-Lira 2000:37 on other cues for interrogativity).

Regarding the historical origin of the Caribbean Spanish circumflex interrogative contour, Quilis (1987:128) makes two interesting observations. First, he notes that the same pattern is also found in Canary Island Spanish, a dialect with which Caribbean varieties share many other features (see also Dorta 2000). One may thus reasonably suspect that this interrogative pattern spread from the Canary Islands to the Caribbean. Second, Quilis indicates that a somewhat similar contour is also found in standard European Spanish in what he calls 'preguntas relativas,' which is basically the same concept as 'checks' in the terminology used by Grice and Savino (2003). It is thus at least possible that what originally was a check contour was reinterpreted in the Canary Islands and the Caribbean as a query. Given the widespread use of circumflex intonation in checks, the reinterpretation of this contour as the unmarked yes/no question pattern independently in several areas seems more likely than a common origin for this interrogative contour in all Romance varieties that have some version of it.

### 3. *Summary*

The study of intonation has become a very active area of research in recent years. As our knowledge of the intonational details of the Romance languages increases, it seems sensible to start to consider whether the facts can be examined from a comparative and diachronic perspective. Historical

'intonology' presents specific challenges, as I have noted, but, with all necessary caveats in mind, progress in this area does not seem entirely out of the question. In this chapter I have also explored some basic features of the intonational system of Occitan, a key language for these purposes, from a pan-Romance point of view, going somewhat beyond what I was able to advance in Hualde (2003).

#### REFERENCES

- Arvaniti, Amalia, D. Robert Ladd, & Ineke Mennen. 2000. "What Is a Starred Tone? Evidence from Greek". *Papers in Laboratory Phonology V: Acquisition and the lexicon* ed. by Michael Broe & Janet Pierrehumbert, 119-131. Cambridge: Cambridge University Press.
- Avesani, Cinzia. 1995. "ToBI: un sistema di trascrizione per l'intonazione italiana". *Metodologie di analisi e di descrizione delle caratteristiche prosodiche e intonative dell'italiano (Atti delle 5e Giornate di Studio del Gruppo di Fonetica Sperimentale dell'A. I. A.)* ed. by Gianni Lazzari, 85-98. Rome: Associazione Italiana di Acustica.
- Baldi, Philip. 1999. *The Foundations of Latin*. Berlin: de Gruyter.
- Beckman, Mary, Manuel Díaz-Campos, Julia Tevis McGory, & Terrell A. Morgan. 2002. "Intonation Across Spanish, in the Tones and Break Indices Framework". *Probus* 14.9-36.
- Cid Uribe, Miriam & Héctor Ortiz-Lira. 2000. "La prosodia de las preguntas indagativas y no-indagativas del español culto de Santiago de Chile". *Lingüística Española Actual* 22.23-49.
- de Moraes, João Antônio. 1998. "Intonation in Brazilian Portuguese". *Intonation Systems* ed. by Daniel J. Hirst & Albert Di Cristo, 179-194. Cambridge: Cambridge University Press.
- Di Cristo, Albert. 1998. "Intonation in French". *Intonation Systems* ed. by Daniel J. Hirst & Albert Di Cristo, 195-218. Cambridge: Cambridge University Press.
- D'Imperio, Mariapaola. 2002. "Italian Intonation: An overview and some questions". *Probus* 14.37-69.
- Dorta, Josefa. 2000. "Entonación hispánica: interrogativas no pronominales vs. Pronominales". *Lingüística Española Actual* 22.51-76.
- Elordieta, Gorka. 2003. "The Spanish Intonation of Speakers of a Basque Pitch-Accent Dialect". *Catalan Journal of Linguistics* 2.67-95.
- Elordieta, Gorka & José I. Hualde. 2003. "Tonal and Durational Correlates of Accent in Contexts of Downstep in Lekeitio Basque". *Journal of the International Phonetic Association* 33.195-209.
- Estebas-Vilaplana, Eva. 2003. "The Modelling of Prenuclear Accents in Central Catalan Declaratives". *Catalan Journal of Linguistics* 2.97-114.
- Face, Timothy. 2002. *Intonational Marking of Contrastive Focus in Madrid Spanish*. Munich: Lincom Europa.
- Frota, Sónia. 2002. "Nuclear Falls and Rises in European Portuguese: A phonological analysis of declarative and question intonation". *Probus* 14.113-146.
- Garrido, Juan M. 1991. *Modelización de patrones melódicos del español para la síntesis y el reconocimiento de habla*. Ph.D. dissertation, Universitat Autònoma de Barcelona.
- Grice, Martine. 1995. *The Intonation of Interrogation in Palermo Italian: Implications for intonation theory*. Tübingen: Niemeyer.

- Grice, Martine, D. Robert Ladd, & Amalia Arvaniti. 2000. "On the Place of Phrase Accents in Intonational Phonology". *Phonology* 17.143-185.
- Grice, Martine & Michelina Savino. 2003. "Map Tasks in Italian: Asking questions about given, accessible and new information". *Catalan Journal of Linguistics* 2.153-180.
- Gussenhoven, Carlos. 2000. "On the Origin and Development of the Central Franconian Tone Contrast". *Analogy, Levelling and Markedness: Principles of change in phonology and morphology* ed. by Aditi Lahiri, 215-260. Berlin: de Gruyter.
- Henríquez Ureña, Pedro. 1938. "Mutaciones articulatorias en el habla popular". *El español en Méjico, los Estados Unidos y la América Central* (= *Biblioteca de Dialectología Hispanoamericana* 4) ed. by Pedro Henríquez Ureña, 329-379. Buenos Aires: Universidad de Buenos Aires.
- Hualde, José I. 2003. "Remarks on the Diachronic Reconstruction of Intonational Patterns in Romance With Special Attention to Occitan as a Bridge Language". *Catalan Journal of Linguistics* 2.181-205.
- Hualde José I., Gorka Elordieta, Iñaki Gaminde, & Rajka Smiljanic. 2002. "From Pitch Accent to Stress Accent in Basque". *Laboratory Phonology 7* ed. by Carlos Gussenhoven & Natasha Warner, 547-584. Berlin: de Gruyter.
- Jun, Sun-Ah & Cécile Fougeron. 1995. "The Accentual Phrase and the Prosodic Structure of French". *Proceedings of the 13th International Congress of Phonetic Sciences* ed. by K. Elenius & P. Branderud, vol. II, 722-725. Stockholm: Stockholm University.
- , 2000. "A Phonological Model of French Intonation". *Intonation: Analysis, modelling and technology* ed. by Antonio Botinis, 209-242. Dordrecht: Kluwer.
- , 2002. "Realizations of Accentual Phrase in French Intonation". *Probus* 14.147-172.
- Kaisse, Ellen. 2001. "The Long Fall: An intonational melody of Argentinian Spanish". *Features and Interfaces in Romance: Essays in honor of Heles Contreras* ed. by Julia Herschensohn, Enrique Mallén, & Karen Zagona, 147-160. Amsterdam & Philadelphia: John Benjamins.
- Ladd, D. Robert. 1990. Review of *Intonation: Emotion vs. grammar*, by Dwight Bolinger. *Language* 66.806-816.
- , 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- Loddo, Daniel. 1993. *Contes e racontes dels Monts de La Cauna (Contes des montagnes du Tarn)*. *Anthologie du conte populaire occitan en Midi-Pyrénées*, 2. Gaillac: GEMP/La Talvera. Book and audiocassette.
- Meisenburg, Trudel. 2001. "À propos des caractéristiques prosodiques de l'occitan". *Le rayonnement de la civilisation occitane à l'aube d'un nouveau millénaire (Actes du 6e Congrès International de l'Association Internationale d'Études Occitanes, 12-19 Septembre 1999)* ed. by Georg Kremnitz, Barbara Czernilofsky, Peter Cichon, & Robert Tanzmeister, 553-560. Vienna: Praesens.
- Oakes, Leigh. 2002. "Phonostylistique des annonceurs de la radio: Étude prosodique de textes radiophoniques". *French Language Studies* 12.279-306.
- Ortiz-Lira, Héctor. 1994. *A Contrastive Analysis of English and Spanish Accentuation*. Ph.D. dissertation, University of Manchester.
- , 2000. "La acentuación contextual en español". *Onomazein* 5.11-44.
- Pei, Mario. 1976. *The Story of Latin and the Romance Languages*. New York: Harper & Row.
- Penny, Ralph. 1991. *A History of the Spanish Language*. Cambridge: Cambridge University Press.
- Penson, Roger. 1993. "Accent and Metre in French". *French Language Studies* 3.19-37.
- Posner, Rebecca. 1996. *The Romance Languages*. Cambridge: Cambridge University Press.

- Post, Brechtje. 1999. "Restructured Phonological Phrases in French: Evidence from clash resolution". *Linguistics* 37.41-63.
- . 2000. *Tonal and Phrasal Structures in French Intonation*. Ph.D. dissertation, Katholieke Universiteit, Nijmegen (LOT Series 34). The Hague: Thesus.
- . 2002. "French Tonal Structures". *Proceedings of Speech Prosody 2002* ed. by Bernard Bel & Isabel Marlien, 583-586. Aix-en-Provence: Laboratoire Parole et Langage.
- Prieto, Pilar. 2002a. "Entonació". *Gramàtica del català contemporani* ed. by Joan Solà, Maria-Rosa Lloret, Joan Mascaró, & Mauel Pérez Saldanya, 393-462. Barcelona: Empúries.
- . 2002b. *Entonació: Models, teoria, mètodes*. Barcelona: Ariel.
- Quilis, Antonio. 1987. "Entonación dialectal hispánica". *Actas del I Congreso Internacional sobre el Español de América* ed. by Humberto López Morales & María Vaquero, 117-164. Puerto Rico: Academia Puertorriqueña de la Lengua Española.
- . 1988. "Estudio comparativo entre la entonación portuguesa (de Brasil) y la española". *Revista de Filología Española* 68.33-65.
- . 1993. *Tratado de fonología y fonética españolas*. Madrid: Gredos.
- Riad, Tomas. 1998. "The Origin of Scandinavian Tone Accents". *Diachronica* 15.63-98.
- . 2000. "The Origin of Danish *Stød*". *Analogy, Levelling and Markedness: Principles of change in phonology and morphology* ed. by Aditi Lahiri, 261-300. Berlin: de Gruyter.
- . 2003. "Diachrony of the Scandinavian Accent Typology". *Diachronic Prosody* ed. by Paula Fikkert & Haike Jacobs, 91-144. Berlin: de Gruyter.
- Smiljanic, Rajka. 2002. *Lexical, Pragmatic and Positional Effects on Prosody in Two Dialects of Croatian and Serbian: An acoustic study*. Ph.D. dissertation, University of Illinois at Urbana-Champaign.
- Sosa, Juan Manuel. 1999. *La entonación del español*. Madrid: Cátedra.
- Truckenbrodt, Hubert. 2002. "Upstep and Embedded Register Levels". *Phonology* 19.77-120.
- Vallduví, Enric. 1991. "The Role of Plasticity in the Association of Focus and Prominence". *Proceedings of the Eastern States Conference on Linguistics* 7.295-306.
- . 1992. *The Informational Component*. New York: Garland.
- Willis, Erik. 2003. *The Intonational System of Dominican Spanish: Findings and analysis*. Ph.D. dissertation, University of Illinois at Urbana-Champaign.



# ROMANCE SE AS AN ASPECTUAL ELEMENT

PAULA KEMPCHINSKY

*University of Iowa*

## 0. *Introduction*

In ongoing research, I have been developing an analysis of the relationship between argument structure and aspectual structure in terms of the linking of arguments of a predicate to temporal subevents of the predicate, following closely work by Zagona (1999). This approach is based on a Pustejovskian view of event structure (Pustejovsky 1991), according to which a given predicate may have at most two subevents, which Zagona labels E1 (informally, ‘INITIATE’) and E2 (approximately, ‘TRANSITION/RESULT’).<sup>1</sup> In Kempchinsky (2000) I proposed that arguments link to temporal subevents in certain specified functional projections, which I labeled ‘Event Phrases,’ where the Event Phrase in which linking to E2 takes place is an aspectual phrase located between the two layers of the verb phrase, following Travis (2000), and the Event Phrase in which linking to E1 occurs is a functional projection just above vP, similar to the Event Phrase proposed in that position by Travis. I refer to these projections here as AspP and EP, respectively, keeping in mind that they are both functional projections with aspectual content. Contra Travis, however, I assume that vP rather than AspP is the locus of object Case checking; that is, the light v carries the relevant set of [phi-features,Case] to enter into a checking relation with an object DP (Raposo & Uriagereka 1996). Following general principles of the Minimalist Program, movement of a DP to Spec of the relevant aspectual head cannot be forced for interpretive reasons; however, the DP can move through Spec of the relevant Event Phrase on its way to its Case-checking position. In (1) I illustrate how the overall system works with an accomplishment predicate:

- (1) a. *La niña comió el helado.*  
“The little girl ate the ice cream.”  
b. [TP *la niña comió*+T [EP (*la niña*) t<sub>v+v</sub> [VP *el helado* ]]]  
[VP (*la niña*) t<sub>v+v</sub> [AspP (*el helado*) t<sub>v</sub> [VP (*el helado*) t<sub>v</sub> ]]]

---

<sup>1</sup> This division into at most two subevents differs from a view of event structure according to which accomplishment predicates, for example, are composed of three subevents, INITIATE (or CAUSE), PROCESS, and CHANGE OF STATE, as in Folli (2002), among others; see section 3.1.



Example (1b) assumes overt object Case checking, but the results will be similar with covert checking. On the assumption that the temporal subevents of the V are part of its interpretative features, these features will become sublabels of T, along with the relevant uninterpretable features for Case checking.

A logical testing ground for this framework is the analysis of sentences with the reflexive clitic SE (Spanish/French/Portuguese *se*, Italian *si*). There have been numerous attempts to provide a unified analysis of the various syntactic manifestations of SE. Many of these analyses converge on a view of SE as a valency-reducing morpheme, whereby it suspends or absorbs a thematic argument of the verb and/or some structural Case (cf. Burzio 1986; Wehrli 1986; Manzini 1986; Cinque 1988; Grimshaw 1990, among others). In a more recent analysis, McGinnis (1998) analyzes SE as a non-DP external argument of the verb, that is, with first merge in Spec,vP, while for Folli (2002) SE is a verbal operator in v (for SE as an operator also see Zubizarreta 1987).

I propose that SE is primarily an aspectual element that either links to, introduces, or absorbs a temporal subevent. In some constructions, this will be E2, so that SE may be interpreted as telic; however, to say that SE is an aspectual element is not to say that SE is necessarily always a marker of telicity. In some cases, the linking of SE to a temporal subevent will result in a reduction of valency because SE will thus deprive an argument of a temporal subevent to link to; in other cases, SE will actually add event structure. The specific ways in which SE interacts with event structure will yield as a consequence the variety of syntactic constructions in which it is found.

My goal here is to defend this analysis of SE in relative detail for two manifestations of SE: reflexive SE and ergative or inchoative SE. In other work (Kempchinsky 2003) I extend the analysis to middle SE and passive SE. The chapter is organized as follows. In section 1, I present my working assumptions on the morphosyntactic nature of the morpheme SE and outline in greater detail the aspectual framework. In section 2, I present the analysis of reflexive SE, and in section 3 the analysis of ergative/inchoative SE, focusing in particular in section 3.2 on the nature of the upper Event Phrase. Finally, in section 4, I explore some implications of the analysis.

## 1. *Framework of the analysis*

### 1.1 *The nature of the morpheme SE*

Of the major Romance languages, Spanish and Italian show the widest range of constructions with SE; an approximate classification with examples from Spanish appears in (2):<sup>2</sup>

- (2) Nonparadigmatic (only third-person verbal forms):
- a. Impersonal: *Se trabaja duro en el mundo académico.*  
“One works hard in the academic world.”
  - b. Passive: *Se construyeron (varias) casas allí.*  
“(Various) houses were constructed there.”
  - c. Middle: *Las camisas de algodón se lavan fácilmente.*  
“Cotton shirts wash easily.”
- Paradigmatic
- d. Ergative: *Las ventanas se rompieron durante la tormenta.*  
“The windows broke during the storm.” (inchoative)
  - e. Reflexive: *La niña se miraba (a sí misma) en el espejo.*  
“The little girl observed herself in the mirror.”
  - f. Causative: *Juan se afeita en la barbería (para impresionar a sus amigos).*  
“Juan gets shaved at the barbershop (in order to impress his friends).”
  - g. Inherent: *Los estudiantes siempre se quejan de las clases.*  
“The students always complain about their classes.” (antipassive)
  - h. Aspectual: *El niño se le comió toda la leche a su hermano.*  
“The little boy drank up all the milk on his brother.”

It is usually assumed that French lacks passive SE, at least with perfective tenses (Ruwet 1972); it also appears to lack causative and aspectual SE. European Portuguese, according to informants that I have consulted, also lacks causative SE.

There is general agreement on the idea that SE is phi-defective; Burzio (1991) proposes that SE has no inherent features of its own. In contrast, Kayne (2000) proposes that SE—more specifically, the *s-* morpheme—is the third-person form, parallel with first-person *m-* and second-person *t-*. He argues that SE is defective in lacking number. I adopt this view; thus, the only phi-feature that SE bears is [person]. In this respect, it is similar to English expletive *there*,

<sup>2</sup> In both the generative and the nongenerative literature, there is wide divergence on the terminology and classification of constructions with SE. The labels in (2) are meaningful, in that they distinguish SE constructions with differing syntactic properties, but are not necessarily contentful in and of themselves. I do not include here the third-person indirect-object use of *se* as in *Se lo di a Ana* “I gave it to Ana”; for arguments that this is an instantiation of SE and not merely a morphophonological variant of *le/les*, see Uriagereka (1997).

and we might speculate that the minimal set of phi-features is precisely [person]. Following Raposo and Uriagereka (1996), the fact that SE bears only [person] entails that it cannot check Case on *v*, because *v*, when it bears a Case feature, also has a full set of phi-features that must be checked and deleted.

I further assume that SE as a clitic is a minimal-maximal projection; however, it is not a DP but rather a  $\phi$ P, in the inventory of pronoun types proposed by Déchaine and Wiltschko (2002). They argue that first- and second-person clitics in French, for example, are  $\phi$ Ps, and as such allow for a bound variable reading under ellipsis, as in (3):

- (3) a. *Je pense que la police m'a vu, et Marie le pense aussi.*  
 b.  $\lambda x$  [x thinks that the police saw me] and  $\lambda y$  [y thinks that the police saw me]  
 c.  $\lambda x$  [x thinks that the police saw me] and  $\lambda y$  [y thinks that the police saw y]  
 (Déchaine & Wiltschko 2002:431)

In contrast, English first- and second-person pronouns, as DPs, do not allow the reading in (3b). Crucially, SE does allow a bound variable reading, as in (4):

- (4) a. *Ana piensa que se defiende bien en italiano, y Luis también lo piensa.*  
 "Ana thinks that she gets by well in Italian, and Luis also thinks so."  
 b.  $\lambda x$  [x thinks that x gets by well in Italian] and  $\lambda y$  [y thinks that y gets by well in Italian]

Déchaine and Wiltschko propose that  $\phi$ Ps may be either predicates or arguments, while DPs may only be arguments. As a clitic, SE may merge either into a head position or a Spec position. I propose that these two possibilities correspond to its dual nature as a  $\phi$ P: When initial merge is in head position, SE is a predicate; when initial merge is in Spec position, SE functions as an argument.

### 1.2 *The syntax of (lexical) aspect*

Following Travis (2002), I assume that within *v*P there are three potential heads that enter into the determination of lexical aspect: the head X of XP (e.g., PP or AP) in VP, the head Asp of AspP between VP and *v*P, and the head *v* of *v*P, as in (5):

## (5) L-syntax aspectual heads (Travis 2002)

v	Asp	X
Functor category	Functional category	Lexical category (AP/PP)
• restricted class	• closed class	• open class
• natural endpoint beginning point	• natural endpoint beginning point	• natural endpoint arbitrary point

As a lexical category, an aspectual head X will generally be realized overtly (cf. English *hammer the nail flat*), while the functional category Asp and the functor category v may or may not have overt morphology. In a language such as Spanish, these heads will have only abstract features to be checked by the verb as it raises to T, as seen in (1b).

In contrast, the head of the upper Event Phrase in (1) is an S-syntax aspectual head, distinct from v (see section 3.2).<sup>3</sup> Adopting this general syntactic approach to aspect, I propose that SE as a  $\phi$ P may appear either in an aspectual head position or in the Spec of an aspectual position, yielding the preliminary aspectual classification of SE in (6):

- (6) SE as an aspectual head:
- a. Inchoative SE: head of AspP, introduces CHANGE OF STATE
  - b. Middle SE: head of vP, suspends INITIATE
- SE as an element in Spec of an aspectual projection:
- c. Reflexive SE: in Spec,Asp, links to E2 (TRANSITION/RESULT)
  - d. Passive SE: in Spec,vP, links to E1 (INITIATE)

Under this analysis, then, reflexive SE and ergative/inchoative SE are instantiations of SE that are operative in the ‘lower’ aspectual level of AspP.

## 2. Reflexive SE

### 2.1 Properties of SE in reflexive constructions

There are two particular properties of constructions with reflexive SE that need to be explained. First, although reflexive SE appears with verbs that are otherwise transitive, it is well known that the reflexive clitic has an intransitivizing effect, as shown by Case alternations under causative *faire* in French (Kayne 1975). Second, reflexive SE in Spanish allows clitic doubling, regardless of the lexical aspectual class, as in (7):

<sup>3</sup> The l-syntax/s-syntax distinction originates with Hale and Keyser (1993), who assume that l-syntax is ‘syntax in the lexicon.’ Travis (2000) argues that both ‘l-syntax’ and ‘s-syntax’ are in the syntactic component; what the l-syntax defines is something like ‘possible word.’

- (7) a. *Ana se conoce bien (a sí misma)*. (stative)  
 “Ana knows herself well.”  
 b. *Ana se observaba (a sí misma) en el espejo*. (activity)  
 “Ana observed herself in the mirror.”  
 c. *Ana se transformó (a sí misma) de niña maleducada a mujer sofisticada*.  
 (accomplishment)  
 “Ana transformed herself from an ill-behaved child to a sophisticated woman.”  
 d. *Ana se reconoció (a sí misma) en la foto*. (achievement)  
 “Ana recognized herself in the photo.”

However, doubling is not always acceptable with SE with ‘verbs of grooming’ as in (8) (I return to this issue in section 2.2):

- (8) a. (??) *El niño se viste a sí mismo*.  
 “The child dresses himself.”  
 b. *El niño ya se viste a sí mismo*.  
 “The child already dresses himself.”  
 c. *Juan se viste elegantemente (??a sí mismo)*.  
 “Juan dresses (??himself) elegantly.”  
 d. *Juan se afeita con una maquinilla eléctrica (??a sí mismo)*.  
 “Juan shaves (??himself) with an electric razor.”  
 (8c,d from Otero 1999)

The acceptability of doubling of *se* correlates negatively with the acceptability of omission of *se* in Spanish under causative *hacer*: Such omission is optional in cases that resist doubling (with some variability across speakers, indicated with ‘%’), but impossible in cases that fully allow doubling,<sup>4</sup> as in (9):

- (9) a. *La madre hizo bañar (%se) al niño*.  
 “The mother made the child bathe.”  
 b. *Su novia hizo afeitarse (%se) a Juan*.  
 “His girlfriend made Juan shave.”  
 c. *La profesora hizo criticarse a los estudiantes*.  
 “The professor made the students criticize themselves.”(reciprocal interpretation also possible)  
 c.′ *La profesora hizo criticar a los estudiantes*.  
 “The professor had (someone) criticize the students.”  
 \*“The professor made the students criticize themselves/each other.”  
 d. *La madre le hizo mirarse/\*mirar en el espejo*.  
 “The mother made him look at himself in the mirror.”

<sup>4</sup> This contrasts both with Italian, which systematically disallows the clitic *si* under *fare* (Burzio 1986), and with French, where *se* is obligatory under *faire* as a reflexive, but optional in the ergative construction (Ruwet 1972). See section 4.

As noted in the introduction, sentences with reflexive SE are not automatically telic. In an object-delimitation language such as Spanish, both the temporal structure of the verb—its subevents—and the syntactic and semantic characterization of the object enter into the aspectual calculus. Thus, stative and activity predicates, lacking the subevent E2 (TRANSITION/RESULT), will remain atelic in the presence of reflexive SE. SE does, however, fulfill the role of a quantized/specific object in the determination of telicity with accomplishment and transitive achievement predicates, as shown in (10):

- (10) a. *Ana describió problemas generales \*en diez minutos.*  
 “Ana described general problems \*in ten minutes.”  
 b. *Ana describió los problemas fundamentales en diez minutos.*  
 “Ana described the fundamental problems in ten minutes.”  
 c. *Ana se describió al psiquiatra en diez minutos.*  
 “Ana described herself to the psychiatrist in ten minutes.”

Because the temporal structure of the verb is crucial for the determination of telicity, part of the computation of aspectual meaning must involve movement of the V through the head of Asp. I assume that Asp, as a functional category, has an uninterpretable temporal feature; in contrast, as mentioned in the introduction, the temporal subevents of a predicate are interpretable features on the verb. Thus, movement of the V through Asp checks and deletes the uninterpretable features on the functional head.

Consider then the derivation of a sentence with reflexive SE in French or Italian, languages in which doubling of the reflexive clitic is not possible. SE is merged directly into Spec,Asp; thus the V may move through the head of Asp as necessary. In fact, it must do so. The structure by the time that the subject is merged into Spec,v is shown in the Italian example (11):

- (11) a. *Gianni si guarda nell' specchio.*  
 “Gianni looks at himself in the mirror.”  
 b. [<sub>VP</sub> Gianni guarda [<sub>AspP</sub> si t<sub>V</sub> [<sub>VP</sub> t<sub>V</sub> nell' specchio ]]]

*Si* (SE) in Spec,Asp receives a temporal role directly in that position, following Ritter and Rosen (1998). They propose that an argument may be directly inserted into Spec,FP (where FP is the functional projection for delimiting events, equivalent to our AspP), and its interpretation is thus determined by its event role. But recall that SE, because it is phi-defective, cannot check Case on v. At the same time, another DP potentially merged in the lower VP would be blocked from checking its Case by the presence of SE in Spec,Asp—hence, the intransitivizing effect of SE. The only way for the derivation to converge is for

v to not enter the numeration with uninterpretable features that need to be checked and deleted.

Now, reflexive SE is paradigmatic, and so the account must extend to first- and second-person reflexives. But of course SE is also a clitic, and as such will need to cliticize to T. It is the cliticization of SE that will eventually allow it to inherit specific person and number features from its antecedent, at the end of the derivation, when SE is cliticized to T and the subject DP is in Spec,T, as in Raposo and Uriagereka (1996).<sup>5</sup>

## 2.2 Doubling of SE in Spanish and verbs of personal grooming

The examples of reflexive SE given in the literature are often constructions with what Saltarelli (1994) termed ‘verbs of personal grooming,’ such as *vestirse* ‘to get dressed/to dress oneself’ and *afeitarse* ‘to shave (oneself)’ in (8). Doubling of *se* with these verbs varies in acceptability from speaker to speaker, and as seen in (8), this variation is influenced by the presence of adverbial elements in the sentence.

Torrego (1995) argues that the SE of doubled constructions is distinct from other instances of SE, and is actually the head of a DP projection. She dubs this instance of SE ‘expletive’ SE, and shows, on the basis of raising out of conjuncts, that SE and *sí mismo* must start as a syntactic unit. Following the general template for doubled clitics in Spanish proposed by Uriagereka (1997), I assume the structure in (12) for doubled SE:

- (12) a. [XP DOUBLE [ CLITIC [AgrP pro Agr [SC t<sub>DOUBLE</sub> t<sub>pro</sub> ]]]]  
 b. [XP (a) *sí* [ se [AgrP [ *mismo* [pro]] Agr [SC t<sub>sí</sub> t<sub>[mismo][pro]] ]]]]<sup>6</sup></sub>

This DP, as a normal DP argument, initially merges into canonical direct-object position, and like any other direct-object DP, may move to Spec,Asp to link to the appropriate temporal role. That is to say, these are full transitive structures, hence the impossibility of omission of *se* under *hacer*. Given the correlation noted earlier between the possibility of doubling and the ungrammaticality of omission of *se*, it must be the case, then, that *se* with a

<sup>5</sup> Alternatively, SE as a maximal projection may move to an outer Spec position of vP, a move that may be motivated by the need for SE to get to the edge of the phrase. In the final section, I suggest a different intermediary move, with SE in the head of the upper Event Phrase.

<sup>6</sup> Torrego (1998) proposes that the presence of the prepositional dative *a* is related to overt movement of the double (in this case, *sí mismo*) to the outer Spec of v. She revises this in Torrego (2002), where she claims that the *a*-phrase moves overtly to Spec,Asp sandwiched between the two verbal layers. Either account is compatible with the analysis here.

transitive verb in Spanish always merges as a DP in canonical object position, with either overt doubled *a sí mismo* or nonovert [pro].

Now this in principle means that verbs of grooming with *se* in Spanish have two possible derivations, one with SE as  $\phi$ P in Spec,Asp and the other with SE as DP in canonical object position, with individual variation across speakers with respect to which verbs they allow with what we might informally term a ‘double subcategorization.’ In the derivation with SE in Spec,Asp, as in the derivation of Italian examples such as (11), the *v* must enter the numeration without uninterpretable features to check, since SE is defective. Conversely, when SE merges as part of a complex DP, *v* will enter into a checking relation with the double ([pro] or *sí mismo*). Such ‘double subcategorization’ is exactly parallel to the situation that Saltarelli (1994) outlines for Latin. The chart in (13) is adapted from his work, with the last column, for Spanish, being my addition:<sup>7</sup>

(13)

Latin	Italian	English	Spanish
<i>se lauat</i>	<i>lava sé stesso</i>	washes himself	<i>se lava (a sí mismo)</i>
<i>lauatur</i> (PASS)	<i>si lava</i>	is washing/washes	<i>se lava (*a sí mismo)</i>
<i>se accingit</i>	<i>veste sé stesso</i>	dresses himself	<i>se viste (a sí mismo)</i>
<i>accingitur</i> (PASS)	<i>si veste</i>	is dressing/dresses	<i>se viste (*a sí mismo)</i>

What the chart in (13) shows is that the ‘detransitivizing’ use of SE in Romance corresponds to the synthetic passive voice in Latin, as Saltarelli pointed out. I return to this point in section 3.2.

### 3. *Ergative/inchoative SE*

#### 3.1 *Ergative/inchoative SE as a marker of telicity*

The SE of ergative pairs (in Burzio’s (1986) terminology) has been extensively studied; typical examples are those of (14).

- (14) a. *El viento rompió los cristales.*  
 “The wind broke the glass panes.”

<sup>7</sup> A reviewer asks whether the possibility of two structures has any aspectual consequences, noting that undoubled *Juan se viste elegantemente* “Juan dresses elegantly” seems to focus on the result rather than the process. I agree with the observation, but I am not sure that it is a result of the structure (compare *Juan se viste rápidamente* “Juan dresses rapidly”). As far as I have been able to determine thus far, doubled and nondoubled verbs of grooming have the same range of possible aspectual interpretations, although the nondoubled structure strongly favors the result reading (and similarly for the English nonreflexive counterpart).



- b. *Los cristales se rompieron (por sí solos)/\*por el viento/\*para quitarlos del marco más fácilmente.*  
 “The glass panes broke (by themselves)/\*by the wind/\*in order to remove them from the frame more easily.”
- c. *El ruido de la calle me despertó.*  
 “The noise from the street woke me.”
- d. *(Yo) me desperté.*  
 “I awoke.”

In Spanish, ergative SE may not be doubled, and it may (preferentially, for some speakers) be omitted under *hacer*, as in (15) and (16):

- (15) a. *Los cristales se rompieron (\*a sí mismos).*  
 “The glass panes broke.”
- b. *Yo me desperté a mí misma.*  
 “I woke myself.”/\*“I awoke.”
- (16) a. *El viento hizo romper(se) los cristales.*  
 “The wind made the glass panes break.”
- b. *El ruido hizo despertar(se) al niño.*  
 “The noise made the child awaken.”

The prevailing view of ergative SE in syntactic theory has been that in these constructions the reflexive clitic absorbs (or, for Cinque 1988, ‘suspends’) the external  $\theta$ -role and the internal accusative Case. However, as pointed out by Folli (2002), the suppression of the external  $\theta$ -role is independent of the presence of SE, since there are transitive/inchoative alternations without SE, as in the Italian examples of (17):

- (17) a. *Gianni ha affondato la barca.*  
 “Gianni has sunk the ship.”
- b. *La barca è affondata.*  
 “The ship has sunk.”

Folli encodes the possibility for a verb to appear in a transitive or inchoative structure in terms of verbal features, whereby an alternating verb such as *affondare* will bear the features [(+v), +V]. What SE expresses in the inchoative construction is completion of the action, that is, telicity. Thus, verbs that are necessarily telic obligatorily appear with *si* in the inchoative, verbs that are necessarily atelic may not appear with *si*, and verbs that may or may not be telic appear optionally with *si*, as in (18):<sup>8</sup>

<sup>8</sup> Folli (2002) notes that within Italian there are dialectal differences with respect to which verbs fall into which class; this is also true across languages (cf. Spanish *Las patatas \*(se) cocieron* “The potatoes cooked”).

- (18) a. *La finestra \*(si) è rotta.*  
 “The window broke.”  
 b. *La temperatura \*(si) è cambiata.*  
 “The temperature changed.”  
 c. *Le patate (si) sono cotte.*  
 “The potatoes cooked.” (Folli 2002:92)

Tellingly, with the verbs in the optional *si* class, when *si* appears, the only interpretation of the sentence is telic, as shown in (19) (Folli 2002:128); similar examples in Spanish from Bruhn de Garavito (2002) are in (20):

- (19) a. *Il cioccolato è fuso per pochi secondi/in pochi secondi.*  
 “The chocolate melted for a few seconds/in a few seconds.”  
 b. *Il cioccolato si è fuso \*per pochi secondi/in pochi secondi.*  
 “The chocolate melted \*for a few seconds/in a few seconds.”  
 c. *La casa è bruciata (per un’ora), ma non è bruciata.*  
 “The house burned (for an hour), but it isn’t burnt.”  
 d. *La casa si è bruciata, \*ma non è bruciata.*  
 “The house burned, \*but it isn’t burnt.”  
 (20) a. *El agua hirvió.*  
 “The water boiled.”  
 b. *El agua se hirvió toda.*  
 “The water all boiled.”  
 c. *Esa madera \*(se) quemó durante más de una hora.*  
 “That wood burned for more than an hour.”  
 d. *Esa madera \*(se) quemó en menos de una hora.*  
 “That wood burned in less than an hour.”

Aspectually, predicates with ergative SE are intransitivized forms of transitive accomplishments. In Folli’s analysis, a full transitive accomplishment has three verbal layers: vP, VP, and RvP, where the head Rv is the predicate encoding CHANGE OF STATE. In contrast, sentences with ergative *si* lack the vP layer, and *si* itself is analyzed as a verbal element in the head of VP. The actual lexical verb merges as the head of RvP and raises to adjoin to *si*. This is partially motivated by the idea that *si* as a +V element is semantically highly underspecified, so it must have a full lexical verb adjoined to it (21):

- (21) [<sub>VP</sub> *la finestra* [<sub>V</sub> *si* ] [<sub>RvP</sub> (*la finestra*) [<sub>Rv'</sub> *chiudere* ...]]]

*Si* as a reflexive operator triggers identification of the two Spec positions; hence the DP *la finestra* raises from Spec,Rv to Spec,V.

What Folli’s (2002) analysis fails to capture, it seems to me, is precisely the correlation between the presence of SE and a result state, since *si* merges as the

head of the process verbal layer. I propose that SE in these sentences, rather than being an argumental element merged into Spec,Asp, is merged directly as the head of Asp, as in (22):

- (22) [<sub>AspP</sub> *la finestra* [<sub>Asp'</sub> *si* [<sub>VP</sub> (*la finestra*) *chiudere* ... ]]]

As a head, SE does not cause a Minimal Link Condition problem for the DP-internal argument of the V to raise to Spec,T to check T's feature set of [phi-features/Case], passing through Spec,Asp on the way. Recall that Asp is the locus of linking E2—TRANSITION/RESULT—to the DP argument. I propose that what happens is that SE, as a predicate, introduces the temporal role (subevent) CHANGE OF STATE. The internal argument DP that moves through Spec,Asp thus links specifically to that type of E2 via Spec-head agreement.

Thus, the structure of an inchoative sentence lacking SE such as (20a) will be like the structure of an inchoative sentence with SE such as (20b) in having only one verbal layer, VP; they will differ by the absence versus presence of AspP, with SE as the head. As noted briefly earlier, it is often argued that ergatives involve a suppression of the external causer, as evidenced by the ungrammaticality of adjunct purpose clauses and agent-oriented adverbs, and this would appear to follow straightforwardly from the lack of vP.

Nevertheless, Higginbotham (1997) observes that inchoative events have a cause, in that there can be a causal explanation of why the event happened; but there is no argument in the semantic representation of the sentence to be taken as the 'individuation of the cause.' Hence, causal-type adjuncts such as *por sí solo* in Spanish or *da sè* in Italian ("by itself") are possible. Folli (2002) in fact shows that *da sé* is licit in an inchoative sentence only if *si* is also present (23).

- (23) a. *La porta si è aperta da sé.*  
 "The door closed by itself."  
 b. \**La temperatura è cambiata da sé.*  
 "The temperature changed by itself."

A parallel phenomenon to the ergative use of an otherwise reflexive pronominal form is found in Kannada, as analyzed by Lidz (2001). This language has a verbal reflexive morpheme that must be present on the verb when the anaphor *tanu* "self" co-occurs as a coargument of its antecedent. In ergative structures, this morpheme is optionally present, as in (24) and (25):

- (24) a. *Hari tann-annu hogaL-i-koND-a.*  
 Hari self-ACC praise-PP-REFL.PAST-3S.M  
 "Hari praised himself."  
 b. \**Hari tann-annu-hogaL-id-a.*

- (25) a. *Baagil-u much-i-tu.*  
 door-NOM close-PAST-3S.N  
 “The door closed.”  
 b. *Baagil-u much-i-koND-itu.*

However, if causal adjuncts are present in such structures, for example, the dative *gaaL-ige* “the wind” in (26), then the verbal reflexive morpheme is obligatory:

- (26) a. *GaaL-ige baagil-u much-i-koND-itu.*  
 wind-DAT door-NOM close-PP-REFL.PAST-3S.N  
 “The door closed because of the wind.”  
 b. \**GaaL-ige baagil-u much-i-tu.*

The licitness of causal adjuncts is somewhat of a mystery if the external  $\theta$ -role is not syntactically projected in the construction. In Lidz’s account, the presence of the verbal reflexive morpheme is evidence of a mismatch between the verb’s thematic and aspectual roles. Thus, in (26), the thematic argument *baagil-u* “the door” is linked to the complex aspectual role [cause [change]] , leaving the aspectual role [act-on] unlinked. This forces us to examine further the aspectual contribution of the higher Event Phrase assumed in the structure given in (1b), and in particular how the temporal subevent of INITIATE is linked in the absence of an external argument. This is the focus of the next section.

### 3.2 *vP, Event Phrase, and Voice Phrase*

Kratzer (1996) argues that the external argument of a verb is not introduced by a lexical element—including here a phonologically null light verb *v*—but rather by a functional head, which in addition checks structural Case on the object (thus preserving, although this is not her intention, Burzio’s generalization). She identifies this functional head as Voice, which she locates just above VP.<sup>9</sup> She further proposes that there is a limited set of (active) Voice heads, one of which adds an agent argument to an action verb, and the other of which adds a ‘holder’ argument to a stative verb. For Kratzer, there is a tight connection between the thematic role of the external argument and the lexical aspectual class of the verb, mediated by the semantic function of Event Identification.

<sup>9</sup> In fact, she does not take a strong position on this matter, noting that in principle Voice can appear anywhere in the hierarchy of inflectional projections as long as it is below Tense, one of whose functions, for her, is to existentially quantify the event argument. In contrast, Saltarelli (1994) proposes that VoiceP is above TP, based on the order of morphemes in Latin synthetic passives.

Thus, this conceptualization of Voice overlaps in many ways with the way in which Travis (2002) views the aspectual properties of the (light verb) *v* head, which she terms a ‘functor’ rather than functional category, following the terminology of Ritter and Rosen (1993). A functor category lacks semantic content and has a closed set of meanings, but potentially more than two, while a functional category can be viewed as having exactly two meanings, expressed in terms of a binary feature. Thus Asp, a functional category, is limited to [ $\pm$ telic], while the meanings of *v* could be limited to DO, BECOME, and CAUSE, as proposed by Folli and Harley (2002).

Nevertheless, Travis also assumes the existence of a functional category external to *v*P, Event Phrase, presumably also coterminous with the upper aspectual Functional Category in Ritter and Rosen (1998). In a similar vein, Zagona (1999) proposes that the syntactic locus of the linking of the subject DP to the relevant temporal role (generally E1, INITIATE) is AgrSP, assumed to be the functional category directly dominating *v*P, an analysis that I explicitly adopted in my earlier account (Kempchinsky 2000). The main thrust of Zagona’s analysis was precisely to link the licensing of DPs as temporal arguments to voice alternations.

The question then is whether the aspectual calculus for the INITIATE subevent involves two separate projections—*v*P and VoiceP/Event Phrase—or only one. I believe that the constructions with ergative SE—and their counterpart with the verbal reflexive morpheme in Kannada—give us evidence in favor of the former position. Recall that *si* in Italian and the verbal reflexive morpheme in Kannada are obligatory in ergative sentences if a causal adjunct is present, as in (23) and (26). Suppose that the ‘limited repertoire’ of Voice, to use Kratzer’s (1996) turn of phrase, is precisely INITIATE (Ritter & Rosen 1998), whereas *v*, following Folli and Harley (2002), assigns the roles of DO, AGENT, or CAUSE. Following the essence of Ritter and Rosen’s analysis, VoiceP—their upper aspectual functional category—is ‘activated’ only if the lower aspectual functional category—here, AspP—is active. What does it mean for AspP to be active? It means that its uninterpretable temporal feature must be valued as +telic, a function fulfilled in these constructions by SE or by the Kannada verbal reflexive morpheme. What the causal adjuncts modify is the subevent of INITIATE; hence these adjuncts are only licit when the relevant telicity-checking morpheme is present in the structure. However, *v* is never present in these ergative sentences, and so strictly agent-oriented adverbials will never be licensed.

Now if the upper Event Phrase is Voice Phrase, and if this projection is activated in telic ergative constructions, then presumably the temporal feature

in this projection must also be checked. Hence it must be the case that the next step of the derivation of sentences with ergative SE, following (21), is (27):

(27) [<sub>vcep</sub> *si* [<sub>AspP</sub> *la finestra* [<sub>Asp'</sub> (*si*) [VP (*la finestra*) *chuidere*]]]]

Saltarelli (1994) shows that the synthetic passive in Latin, like SE, also appeared in ergative structures; thus in Latin the passive affix, as part of the verbal head, checks the temporal features in Asp and Voice.

The obvious question then is why the reflexive clitic SE, the verbal reflexive morpheme in Kannada, and the synthetic passive verbal affix in Latin should all be able to ‘ergativize’ a transitive accomplishment. I offer some speculations on this question in the final section.

#### 4. *Summary and ideas to explore*

In (28) I summarize the essential points of the analysis offered here:

(28) a. Reflexive SE:

- merges as Spec,Asp
- blocks a potential VP-internal object from agreeing with v
- receives a temporal role directly in Spec,Asp
- agrees with antecedent DP in TP

b. Ergative/Inchoative SE:

- merges as head of Asp
- introduces temporal role of CHANGE OF STATE
- VP-internal object links to CHANGE OF STATE in Spec-head relation in AspP
- in absence of vP and hence absence of external argument, SE licenses INITIATE temporal role in head of EP

As noted in footnote 4, *si* in Italian is systematically absent in complements to causative *fare*. For Folli (2002), this follows from her proposal that *si* is in v, on the assumption that *fare*, when present, is in v, taking a VP complement. This account, however, does not straightforwardly extend to either Spanish or French.

There does seem to be some convergence on the idea that causative complements in Italian are essentially monoclausal, while they are either mono- or biclausal in, for example, Spanish and French (cf. e.g., Zubizarreta 1987).<sup>10</sup> Let me therefore sketch briefly here one possible account of the

<sup>10</sup> Davies (1995) argues, based in part on the omission versus appearance of *se*, that causative complements in Spanish and Portuguese evolved from monoclausal to biclausal structures. However, even as early as the 13th century, there are examples of overt *se* in causal complements in Spanish (Davies 2002).

presence versus absence of SE in causative complements. Suppose that in Italian causative *fare* takes a complement that projects only as high as AspP, while in the other languages under consideration the causative complement may project at least as high as VoiceP. Now if SE remains as low as Asp, it can appear under a causative verb—but that is precisely what does not occur, at least in Italian. If it is the case, however, that SE must raise to the head of VoiceP, as in the structure in (27), then SE (i.e., *si*) will not surface in Italian causative complements, but will surface, at least optionally, in the other languages, on the assumption that in these languages the complement is either AspP (i.e., the ‘monoclausal’ structure) or VoiceP.

This is related to the question raised at the end of the previous section, which might be couched as, Why SE? Note that the three cases that I have touched on—SE, the Kannada verbal reflexive morpheme, and the Latin synthetic passive—have two properties in common: They attach (sooner or later in the derivation) to the verb, and they enter into an agreement relation with the surface subject. Now the subject (or its formal features) is ultimately in a checking relationship with Tense, which in turn is in some kind of relationship with the subevents of the verb—either as an existential quantifier or as a temporal anchor, depending on the particular analysis of Tense that one wishes to adopt. If we assume that all temporal roles (subevents) of a predicate must be linked to an argument (an idea already put forth in Pustejovsky 1991), then these morphemes allow one DP argument to be linked to two subevents, while simultaneously linking them to the temporal anchor of the sentence. This role cannot be played by a complex anaphor such as English *himself*, because it is not a verbal affix, and because, as a complex DP, it must have an independent thematic role as well as a temporal role. That, at any rate, is my speculation at this point. Exactly how to formalize this intuition remains a topic of this ongoing research project.

#### REFERENCES

- Burzio, Luigi. 1986. *Italian Syntax: A Government and Binding approach*. Dordrecht: Kluwer.  
 -----, 1991. “The Morphological Basis of Anaphora”. *Journal of Linguistics* 27:1.81-105.
- Bruhn de Garavito, Joyce. 2002. “Is a Subject Always a Subject? The Nature of the NP in Impersonal and Inchoative Constructions in Spanish”. Ms., University of Western Ontario.
- Cinque, Guglielmo. 1988. “On *Si* Constructions and the Theory of Arb”. *Linguistic Inquiry* 19:4.521-581
- Davies, Mark. 1995. “The Evolution of Causative Constructions in Spanish and Portuguese”. *Contemporary Research in Romance Linguistics* ed. by Jon Amastae, Grant Goodall, Mario Montalbetti, & Marianne Phinney, 105-122. Amsterdam & Philadelphia: John Benjamins.

- , 2002. "Modeling Syntactic Change With the Fully Annotated, 100 Million Word 'Corpus del español'". Paper presented at the 6<sup>th</sup> Hispanic Linguistics Symposium, Iowa City, IA, October 2002.
- Déchaïne, Rose-Marie & Martina Wiltschko. 2002. "Decomposing Pronouns". *Linguistic Inquiry* 33:3.409-442.
- Folli, Raffaella. 2002. *Constructing Telicity in English and Italian*. Ph.D. dissertation, University of Oxford.
- Folli, Raffaella & Heidi Harley. 2002. "Consuming Results in Italian and English: Flavors of v". Paper presented at the Workshop on the Syntax, Semantics, and Acquisition of Aspect, Iowa City, IA, May 2002.
- Grimshaw, Jane. 1990. *Argument Structure*. Cambridge, MA: MIT Press.
- Hale, Kenneth & Samuel J. Keyser. 1993. "On Argument Structure and the Lexical Expression of Syntactic Relations". *The View from Building 20* ed. by Kenneth Hale & Samuel J. Keyser, 53-109. Cambridge, MA: MIT Press.
- Higginbotham, James. 1997. "Location and Causation". Ms., University of Oxford.
- Kayne, Richard. 1975. *French Syntax and the Transformational Cycle*. Cambridge, MA: MIT Press.
- , 2000. "Person Morphemes and Reflexives in Italian, French and Related Languages". *Parameters and Universals* ed. by Richard Kayne, 131-184. Oxford: Oxford University Press.
- Kempchinsky, Paula. 2000. "Aspect Projections and Predicate Type". *Hispanic Linguistics at the Turn of the Millenium* ed. by Héctor Campos, Elena Herburger, Alfonso Morales-Front, & Thomas J. Walsh, 171-187. Somerville, MA: Cascadilla Press.
- , 2003. "Aspects of Middle and Passive SE". Paper presented at the 7<sup>th</sup> Hispanic Linguistics Symposium, Albuquerque, NM, October 2003.
- Kratzer, Angelika. 1996. "Severing the External Argument from the Verb". *Phrase Structure and the Lexicon* ed. by Johan Rooryck & Laurie Zaring, 109-137. Dordrecht: Kluwer.
- Lidz, Jeffrey. 2001. "The Argument Structure of Verbal Reflexives". *Natural Language and Linguistic Theory* 19:2.311-353.
- Manzini, Maria Rita. 1986. "On Italian *si*". *Syntax and Semantics: The syntax of pronominal clitics* ed. by Hagit Borer, vol. XIX, 241-262. New York: Academic Press.
- McGinnis, Martha J. 1998. *Locality in A-movement*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Otero, Carlos. 1999. "Pronombres reflexivos y recíprocos". *Gramática descriptiva de la lengua española* ed. by Ignacio Bosque & Violeta Demonte, 1427-1517. Madrid: Espasa-Calpe.
- Pustejovsky, James. 1991. "The Syntax of Event Structure". *Cognition* 4:1-3.47-81.
- Raposo, Eduardo & Juan Uriagereka. 1996. "Indefinite SE". *Natural Language and Linguistic Theory* 14:4.749-810.
- Ritter, Elizabeth & Sara T. Rosen. 1993. "Deriving Causation". *Natural Language and Linguistic Theory* 11:3.519-556.
- , 1998. "Delimiting Events in Syntax". *The Projection of Arguments* ed. by Myriam Butt & Wilhelm Geuder, 135-164. Stanford, CA: CSLI.
- Ruwet, Nicolas. 1972. *Théorie syntactique et syntaxe du français*. Paris: Seuil.
- Saltarelli, Mario. 1994. "Voice in Latin and Romance: On the representation of lexical subjects". *Issues and Theory in Romance Linguistics* ed. by Michael Mazzola, 445-478. Washington, DC: Georgetown University Press.
- Torrego, Esther. 1995. "From Argumental to Non-Argumental Pronouns: Spanish doubled reflexives". *Probus* 7:2.221-241.



- , 1998. *The Dependencies of Objects*. Cambridge, MA: MIT Press.
- , 2002. "Aspect in the Prepositional System of Romance". *Current Issues in Linguistic Theory: Selected proceedings of the 29<sup>th</sup> Linguistic Symposium on Romance Languages* ed. by Teresa Satterfield, Christina Tortora, & Diana Cresti, 337-357. Amsterdam & Philadelphia: John Benjamins.
- Travis, Lisa. 2000. "Event Structure in Syntax". *Events as Grammatical Objects* ed. by Carol Tenny & James Pustejovsky, 145-185. Stanford, CA: CSLI.
- , 2002. "Articulated vPs and the Computation of Aktionsart". Paper presented at the Workshop on the Syntax, Semantics and Acquisition of Aspect, Iowa City, IA, May 2002.
- Uriagereka, Juan. 1997. "Formal and Substantive Elegance in the Minimalist Program". *The Role of Economy Principles in Linguistic Theory* ed. by Chris Wilder, Hans-Martin Gärtner, & Manfred Bierwisch, 170-204. Berlin: Akademie Verlag.
- Wehrli, Eric. 1986. "On Some Properties of French Clitic *se*". *Syntax and Semantics: The Syntax of Pronominal Clitics* ed. by Hagit Borer, vol. XIX, 263-283. New York: Academic Press.
- Zagona, Karen. 1999. "Voice and Aspect". *Grammatical Analyses in Basque and Romance Linguistics* ed. by Jon Franco, Alazne Landa, & Juan Martin, 279-293. Amsterdam & Philadelphia: John Benjamins.
- Zubizarreta, María-Luisa. 1987. *Levels of Representation in the Lexicon and in the Syntax*. Dordrecht: Foris.

## PROTO-ROMANCE \*[w] AND THE VELAR PRETERITES

ERIC LIEF  
Cornell University

### 0. *Introduction*

Old Occitan and Old Catalan, spoken in southern France and northeastern Spain respectively, developed peculiar velar stem reflexes of the Latin perfects in  $-UI$  / $-ui$ /. Examples of these so-called ‘velar preterites’ are given in Table 1.

Latin Perfect /CVV/	Proto-Romance /Cw/	OOc/OCat /g/
DEBUI, DEBUISTI, etc. “owe, ought”	*deβwi, deβwisti, etc.	dek, deγist, etc.
POTUI, POTUISTI, etc. “be able”	*potwi, potwisti, etc.	pok, poγist, etc.
TENUI, TENUISTI, etc. “have”	*tenwi, tenwisti, etc.	tenk, tengist, etc.
MOLUI, MOLUISTI, etc. “grind”	*molwi, molwisti, etc.	molk, moγist, etc.

Table 1: *The velar preterites*

In Romance linguistics, the *communis opinio* has been that these velar preterites are completely the result of a regular sound change  $w > gw$  (e.g., Meyer-Lübke 1895; Anglade 1921; Fouché 1931; De B. Moll [1952] 1991; Lausberg 1966; Klausenburger 1984; Paden 1998). Some of the most frequently cited comments on the development are given in (1).

- (1) a. “Dans les verbes en  $-p$ , l’ $\mu$  s’introduit dans le radical; sinon, comme  $w$  germ., il se transforme en  $g$ , à la finale en  $-c$  [In  $-p$  verbs, the  $\mu$  is introduced into the root; but, like Germanic  $w$ , it changes to  $g$ , or to  $c$  when word final]” (Meyer-Lübke 1895:365).
- b. “Dans les verbes dont la radical n’était pas terminé par un  $p$ , il s’est développé, pour des raisons de phonétique et **non** (my emphasis, EAL) d’analogie, une gutturale entre la consonne finale du radical et la terminaison  $-ui$  [In verbs whose root did not end in  $-p$ , a guttural was introduced, for reasons of phonetics, and not analogy, between the final consonant of the root and the ending  $-ui$ ]” (Anglade 1921:309).
- c. “Sobre la formación  $BW > k, g$ , recuérdense la afinidad fonética entre la bilabial-velar  $w$  y los sonidos plenamente velares  $k, g$  [As for the development  $BW > k, g$ , remember the phonetic similarity of bilabial-velar  $w$  and the velars  $k, g$ ]” (Badía Margarit 1951:318).

- d. “Tots aquests verbs, en passar al català, van sofrir, la velarització de l’element consonàntic +  $\mu$ , i el radical va passar a acabar en  $k$  o  $g$  segons que se’n conservés o no la desinència [All these verbs, on their way to Catalan, underwent velarization of the consonantal element +  $\mu$ , and the root came to end in  $k$  or  $g$ , depending on whether its ending was preserved]” (De B. Moll [1952] 1991:159).
- e. “En cat. y prov. se le antepone una - $g$ - a la - $\mu$ - (\* $awwi$  > \* $agui$ ) [In Catalan and Provençal a - $g$ - was introduced before the - $\mu$ -]” (Lausberg 1966:354).

The common thread of these accounts is that the original /Cw/ sequences developed regularly into a (labial-)velar stop (or fricative) in all environments (i.e.,  $w > gw$ ).

In contrast to the traditional approach, I argue that both sound change (specifically, the development  $bw > gw$ ) and analogy (with the verb “have”) have worked together to create such velar preterites. In the first part of this chapter, I show that there is no support for  $w > gw$  outside of these verb forms. I then reformulate the original sound change (i.e.,  $bw > gw$ ) and propose a mechanism for its analogical extension throughout the Old Occitan and Catalan past-tense systems. The high token frequency of the verb HABUI “have,” which underwent the sound change, is one factor that favored the extension of this velar throughout other verbal paradigms. Finally, to demonstrate the likelihood of analogy with “have,” I examine Spanish and Portuguese perfects where the influence of “have” is unequivocal.

### 1. *The basic reflexes of \*/Cw/ sequences in Western Romance*

While no stops except for the velars /k/ and /g/ freely combine with /w/ in Classical Latin (CL),<sup>1</sup> Proto-Romance developed /Cw/ from original /CVV/ sequences, for example, CL HABUI /ha.bu.i/ “have,” /wi.du.a/ “widow” > \*/habwi/, \*/widwa/. Table 2 shows the reflexes of the original /Cw/ sequences. In this section, I examine the reflexes of these /Cw/ sequences in Western Romance in search of support for the putative sound change  $w > gw$ . After sorting out these data, I show that the only apparent nonverbal examples of  $w > gw$  are of the extremely limited type  $\eta w > \eta gw$ .

#### 1.1 *The labials \*/pw/, \*/bw/*<sup>2</sup>

As Table 3 indicates, the only inherited forms with /pw/ and /bw/ are verbs that formed their perfects with /u/ in Latin. While /pw/ metathesized in probably all of Western Romance (2), /bw/ shows metathesis only in Spanish and Portuguese (3a), but velarization (shaded areas in Table 3) in Occitan,

<sup>1</sup> According to the phonological literature, this was perhaps because /k<sup>w</sup>, g<sup>w</sup>/ were not clusters, but rather complex segments (e.g., Devine & Stevens 1977).

<sup>2</sup> /bw/ came to be realized as [βw] in Romance.

Catalan, and probably French<sup>3</sup> (3b). The different development of these labial sequences is the result of dissimilation of the two labials in inherited \*/bw/, as I show later.

- (2) Metathesis *pw* > *wp*:  
 SAPUI > \*sapwi > \*sawpi > OSp sope “had,” OPg sowbe, OOc sawp, OFr soj, saw<sup>4</sup>
- (3) a. Metathesis *bw* > *wb*: HABUI > \*aβwi > \*awβi > OSp oβe “had,” OPg owve  
 b. Dissimilation (Occitan, Catalan, French) *bw* > *gw*:  
 HABUI > \*aβwi > \*aγwi > OFr oj, aw ~ OOc/OCat ak

	/pw/ SAPUI “know <sub>1PERF</sub> ” *RECIPUI “receive <sub>1PERF</sub> ”	/bw/ HABUI “have <sub>1PERF</sub> ” DEBUISTI “ought <sub>1PERF</sub> ”	/tw/ *STATUALE(M) “statuelike” POTUIT “be able <sub>3PERF</sub> ”	/dw/ VIDUA(M) “widow” *SED(U)IT “be (seated) <sub>3PERF</sub> ”	/nw/ JANUARIU(M), “January” *MINUARE “lessen” (< MINUERE)
OSp	/p/ sope (retsibí)	/b/ oβe (deβí)	/d/ estadal, podo	/bd/, /wd/, /b/ biβða/biwða, soβo	/n/, /ngw/ enero, mingwar
OPg	/wb/ sowbe (retsebí)	/wv/ owve (deví)	/d/ estadal, pode	/wv/, /v/ viwva, seve	/n/, /ngw/ džanejru, mingwar
OCat	/b/ (sabí), (rebí)	/g/ ak, dek	/d/, /g/ estadal, pək	/wb/, /wd/, /g/ viwβa, viwða, sek	/n/, /nv/ džiner, mimjvar
OGc	/b/ sabó (3.PERF)	/g/ ak	/g/ poγosan (PAST.SUBJ)	/bd/, /dw/ beβða, beðwe	Ø, /ng/ džëer, mingar
OOc	/wb/ sawp, retsewp	/g/ ak, dek	/d(o)/, /g/ estad(o)al, pək	/v/, /wd/, /do/, /g/ veva, vewða, veðoa, sek	/n/ dženjer
OFr	/w/, Ø soj, (dial.) saw	/w/, Ø oj, (dial.) aw	/v/, /w/, Ø estavel, poØ, (dial.) powØ	/(d)v/, /w/, Ø ve(ð)ve, (a)syØ	/v/ džävier

Table 2: The reflexes of /Cw/ in Romance

	HABUI “have <sub>1PERF</sub> ”	DEBUI “owe <sub>1PERF</sub> ”	(*)BIB(U)I “know <sub>1PERF</sub> ”	SAPUI “know <sub>1PERF</sub> ”	*RECIPUI “receive <sub>1PERF</sub> ”
OSp	oβe	deβí	beβí	sope	retsibí
OPg	owve	deví	beví	sowbe	retsebí
OCat	ak		bek	sabí	rebí
OGc	ak	part. deγyt		sabó	
OOc	ak, ajk	dek	bek	sawp	recewp
OFr	oj, (dial.) aw	dyj, (dial.) diw	byj, (dial.) 3.PERF) biwt	soj, (dial.) saw	(dial. 3.PERF) recewt

Table 3: \*/pw/ \*/bw/

<sup>3</sup> French is ambiguous here, and, to my knowledge, there is no obstacle in grouping it with Occitan and Catalan.

<sup>4</sup> In Proto-French, \*sawve (with /v/, rather than expected /p/, from leveling, e.g., from infinitive \*savere) develops to either \*saw in the dialects or standard \*(s)wve > \*soe > soj as in \*GRAVA > groe “stony terrain”; see OOc grava “pebble” (Fouché 1931:307).

1.2 *The coronals \*/tw/ and \*/dw/*

There are very few forms continuing \*/tw/ or \*/dw/ that have survived in Romance. The examples given in Table 4 (except for FUTUERE) are all nonverbal.

	FUTUERE “fuck”	MORTUA(M) “dead”	STATUA(M), *STATUALE(M) “statue”	Germanic walja “towel”	VIDUA(M) “widow”
O <sub>Sp</sub>	hoder	mwerta	estadal	toaʒa	biwða
O <sub>Pg</sub>	foder	morta	estadal	toaʎa	viwva
O <sub>Cat</sub>	fotre	morta	estadal	toaʎes (> tovaʎes)	βiwβa, βiwða
O <sub>Gc</sub>		morta			beβða, beðoe
O <sub>Oc</sub>	fotre	morta	estad(o)al	toaʎa	vewða, veðoa
O <sub>Fr</sub>	futrə	mørtə	estavel	tuaʎə	ve(ð)və

Table 4: \*/tw/, \*/dw/

This list becomes smaller and smaller as we eliminate ambiguous forms. First, FUTUERE is problematic since first-person singular FUTUO would have undergone the regular Romance sound change  $w > \emptyset / \_ o$  (cf. COQUERE “cook”: COQUO > \*kəko > It kwəko). In the same vein, MORTUU(M), possibly remade on \*MORTU(M), and QUATT(U)OR are problematic, though loss here would probably have occurred in the complex syllable onset anyway (cf. FEBRA<RIAS> for FEBRUARIAS, Pompeii inscription). In addition, Germanic \*/θwalja/ was probably borrowed into Latin early on as something like \*TUALIA(M), which most likely developed to Proto-Romance \*/toalja/, rather than \*/twalja/, because of the aforementioned constraint on complex syllable onsets. In Table 4, shading indicate these suspect forms.

We are, then, left with two forms: STATUA (and derivatives) and VIDUA. Like MANUALE(M), STATUALE(M) may have been subject to analogy with STATU(M) “position, posture” (Wartburg 1928-). In Gallo-Romance, however, forms like Swiss /etava/ and Wallon /etave/ “fence pole” seem to go back to STATUA(M). Old French /estavel/ and Old Occitan /estad(o)al/,<sup>5</sup> both sharing the meaning “candle of statuelike dimensions, wax,” clearly derive from STATUALE(M). In addition, the similar meaning of O<sub>Sp</sub>/O<sub>Pg</sub>/O<sub>Cat</sub> /estadal/ supports an inherited STATUALE(M).

Holding off on VIDUA for the moment, verb forms like POTUIT (Table 5) seem to support the putative Iberian sound change  $tw > d$ , seen earlier for STATUALE(M). Occitan and Catalan, however, show some velar reflexes for \*/tw/ and \*/dw/. As already mentioned, it is unlikely that analogy played a

<sup>5</sup> O<sub>Oc</sub> *estadoal*, which is similar to [veðoa], perhaps suffered analogy with the learned suffix -UALE(M) (> -oal).

role in the development of STATUALE(M), so the velar outcomes of POTUIT must be analogical, and the shrinking of the shaded area of Table 5 with respect to Table 3 may indicate this fact.

	STET(U)IT “stand <sub>3PERF</sub> ”	POTUIT “be able <sub>3PERF</sub> ”	*CRED(U)IT “believe <sub>3PERF</sub> ”	*SED(U)IT “sit <sub>3PERF</sub> ”
OSp	estjedo	pode	kroβe	soβe
OPg	esteve	pode	kreβe	seβe
OCat	estek	pøk	krek	sek
OGc		pogossan (PRES.SUBJ)	kregud, krezud (PART)	
OOc	estét, esték (very rare)	pøk	krezét	sek
OFr	estyt	pot, (dial.) powt	kryt	(sest)

Table 5: *Derived forms with \*/tw/, \*/dw/*

Better, though not the best, evidence for  $tw > d$  may come from the outcomes of \*BA(T)TUACULU(M), OSp /badaʒo/ and OPg /badalo/ (< Galician /badal, badaʎo/) “clapper (of bell)”; see OOc denominal /bataʎar/ “talk incessantly (like a bell).” Corominas (1980-1991) explains degemination here, supported by Medieval Latin (ML) BATERE for CL BATTUERE “beat,” as the result of confusion with FUTUERE “fuck” (i.e., due to the metaphor “strike” or “beat”). In support of this explanation are the voiceless reflexes OFr/OOc/OCat /fotre/, next to OSp/OPg /foder/. If Ibero-Romance /estadal/ is, in fact, the expected outcome of STATUALE(M), and \*BATUACULU is a good reconstruction, then this instantiates  $tw > d$ .

As for the voiced counterpart \*/dw/, verbs like CREDUIT behave differently from the noun VIDUA(M). However, VIDUA(M) shows so many Romance reflexes that it is extremely challenging to weed out dialectal or later (ML) borrowing and analogy from regular sound change. We find spellings such as OFr *veuve, vedve, vedde, veve*; OOc *veuza, vezoa, veuva, veva*; Old Gascon *bepda, bebda*; Old Bearnese *bedoe, bede*; and OCat *viuva, vilva, viuda, vidua*. As for Ibero-Romance, the most common forms are OPg *viuva*; OLeon *bilda, vilda, vilva*; and OSp *bibda* (Cid), *biuda, bibda, vidua* (Berceo).

Corominas (1980-1991) believes that *viuva* /βiwβa/ is the native Catalan word, while *viuda* /βiwda/ is the outcome of blending with Latinate *vidua*, quite frequent in medieval documents. Like the sequence /nw/ discussed later, Occitan /dw/ may have undergone metathesis in post-tonic environments. The problem is that spellings like OOc *veuza* and *vezoa* could not have been simply borrowed from ML *vidua*, since \**viuza*<sup>6</sup> or \**vizo*a would be the expected

<sup>6</sup> For unattested but expected (reconstructed) forms for any given sound change the symbol \* is used.

learned forms. In this respect, Corominas' account has merit. In (4), analogy or blending of the inherited reflex of VIDUA(M) with the learned outcomes of *vidua* or *viduitas* "widowhood" may explain the actual attested /e/ vocalism of this form, inside and outside of Gallo-Romance, in addition to the linear order of /dw/; see OFr *ve(u)ve* (the ancestor of modern *veuf* ~ *veuve*) ~ *vedve*, OIt *vedova* (unattested before the 14<sup>th</sup> century) instead of \**vedda* ~ *vedovale*, *veduale*, *viduale* (< ML VIDUALE[M]).

- (4) Blending of VIDUA(M) in Romance, example from Old Occitan:  
 VIDUA(M) > *veva* → *veuva*, etc. ← \**viuda* (< ML *vidua*)

If, in fact, the development [dv] > [v], which we find in Iberian verb forms, also holds for Gallo-Romance (6), then attested Old Occitan *veva* and Old French *veve* are easily explained; (5) and (6) show reflexes of /tw/ and /dw/.

- (5) a. /w/-deletion *tw* > *d*:  
 STATUALE(M) > \**estadwal* > OSp/OPg/OOc *estadal*  
 b. /w/-develarization<sup>7</sup> + /t/-deletion *tw* > *v*:  
 STATUALE(M) > \**estadβel* > OFr *estavel*  
 (6) /w/-develarization + /d/-deletion [dw] > [β,v]:  
 VIDUA(M) > \**βedβa* > \**βeβa* → OSp *biwβa*, OPg/OCat *viwva*, OOc *veva*, OFr *veve*

### 1.3 The nasal \*/nw/ <sup>8</sup>

Table 6 presents the major nonverbal sources of /nw/. TENU(M) and JANUA(M) are the only truly underived forms containing \*/nw/ that have come down to us from Latin, and it is very unlikely that analogy with some related base noun has played any role in their development. Although JANUELLA(M) and JANUARIU(M) are both derived from JANUA and JANUS "Italic deity" respectively, it is unlikely that analogy has affected their development in any way, since both JANUA and its derivative contain \*/nw/, and any possible influence of the god JANUS, which has not survived in any daughter language, is highly unlikely. Furthermore, \*MINUARE seems to have diverged from MINUS "less" in all of Romance. I next consider the development of these four forms, as well as Gothic *manwian* for the sake of comparison.

<sup>7</sup> Develarization here represents the common loss of velarity in labial-velars like /w/ or /kw/, for example, *w* > *v/v/β* (cf. Ohala 1979).

<sup>8</sup> Examples with \*/mw/, discounting bimorphemic ones (e.g., /kom-wenire/ "come together"), are rare and probably nonexistent.





maintain loss of the glide before front vowels (i.e., ‘/w/-deletion’) in Ibero-Romance (8a), and epenthesis<sup>9</sup> elsewhere (8b).

- (8) a. /w/-deletion *nw* > *n*:  
 JANUARIU(M)<sup>10</sup> > \*dʒenwejru > \*dʒenejru > OSp *enero*, OPg *dʒanejru*  
 b. /g/-epenthesis *nw* > *ɲgw*:  
 MINUARE > \*menwar > \*miŋgwar > OSp *miŋgwar*, OPg *mī(ŋ)gwar*

Therefore, /nw/ was retained well into Ibero-Romance. Spanish and Portuguese then subsequently deleted the glide independently. In other words, there is no way to sustain something like Proto-Ibero-Romance \*/dʒenella/ in (9) because, had /w/ fallen in the proto-language, the nasal would have deleted in Portuguese.

- (9) \*dʒenella > Pre-Pg \*dʒenela > \*dʒeɛla > OPg \*dʒiɛla ‘window’  
 cf. GENESTA(M) > Pre-Pg \*dʒenesta > \*dʒeɛsta > OPg *dʒiɛsta* ‘broom (tree)’

On this account, \*/w/, when retained (only before /a/), develops to /gw/, as in [mī(ŋ)gwar]. Although possible, the likelihood of independent *nw* > *n* as well as *nw* > *ɲgw* in both Spanish and Portuguese is not too great. As we just saw, however, nasal deletion was blocked in Portuguese, and thus could not have occurred in the proto-language.

The essentially different treatment of the glide depending on the following vowel suggests that the glide had more than one phonetic realization. It is common for labial-velars to lose their velar properties in palatal environments: For example, the CL labial-velar /k<sup>w</sup>/ or /kw/ seems to have developed into a labial palatal (\*[k<sup>ɥ</sup>] or perhaps \*[k<sup>ɥ</sup>]) before a front vowel, as in LAQUEU(M) *lak<sup>w</sup>eu(m)* > \**lak<sup>ɥ</sup>ju* > \**lak<sup>ɥ</sup>ju* > \**latʃju* > It *lattʃo*, Sp *laθo*; and AQUILA(M) *ak<sup>w</sup>ila(m)* > \**ak<sup>ɥ</sup>ila* > \**ak<sup>ɥ</sup>ila* > *ag<sup>ɥ</sup>ila* > Sp/Pg *ágila*. Similarly, I propose the development *nw* > *ny* in palatal environments. This was the state at the Ibero-Romance stage, thus accounting for the blocking of nasal deletion in

<sup>9</sup> Epenthesis here results from coarticulation, that is, intrusive /g/ arises from early closure of the velum in anticipation of the following oral consonant, much like the /p/ found in *something* [sʌmpθɪŋ].

<sup>10</sup> Meyer-Lübke (1890:305-306) argues for the development of unstressed [dʒa] > [dʒe] in all of Romance. If we accept this claim, then the variation in Catalan (*giner* ~ dialectal *janer*) and French (*janvier* ~ dialectal *jenvier*) is probably due to either dissimilation or influence of frequent ML JANUARIUS. Portuguese *janeiro* (cf. Sp *enero*, Cat *giner*) is possibly of learned origin as well. Reflexes of JANUA in French would support Ibero-Romance \*janwa, on which *janela* may have been reconstituted (cf. dialectal *jinela*).

Portuguese (as in Table 7). Later, the glide deleted, probably independently, in both languages.

	*JANUELLA(M)	GENESTA(M)	MINUARE
Proto-Ibero-Romance	*[dʒenwɛlla]	*[dʒenɛsta]	*[minwar]
Epenthesis, Palatalization	*[dʒenʎɛlla]	N/A	*[miŋgwar]
Pre-Portuguese	*[dʒenʎɛlla]	*[dʒenɛsta]	*[miŋgwar]
/n/-deletion	N/A	*[dʒɛɛsta]	N/A
/w/-deletion	*[dʒɛnɛla]	N/A	N/A
OPg	[dʒinɛla]	[dʒiɛsta]	[miŋgwar]

Table 7: *Blocking of nasal deletion in Portuguese*

1.3.2 *\*/nw/ in Ibero-Romance: Catalan.* Like its Ibero-Romance neighbors Spanish and Portuguese, Old Catalan exhibits *nw* > *n* in palatal environments (10a). The development of Gothic \*(AD)MANUIRE > *amanir* is analogous. In contrast, the development of /w/ in nonpalatal environments (10b) differs from that of its Iberian neighbors. Thus Catalan did not undergo /g/-epenthesis. Also note that the raising of unstressed *e* > *i* before the labial-velar (cf. AEQUALE(M) “equal” > \**egwal* > \**igwal* > Sp, Pg, Cat *igwal*) is found across the board in Ibero-Romance.<sup>11</sup>

- (10) a. /w/-deletion (palatal environment) *nw* > *n*:  
 JANUARIU(M) > \**dʒanwɛjru(m)* > *dʒiner*  
 b. /w/-develarization (elsewhere) *w* > *v*:  
 \*MINUARE > \**menware* > *miŋvar*  
 (cf. /gw/ results for OSp and OPg)

1.3.3 *\*/nw/ in Gallo-Romance: Old Gascon.* Unlike the rest of Gallo-Romance, Old Gascon has reflexes of MINUARE (Old Bearnese /*mingar*/ “lesson” [Bayonne, 13<sup>th</sup> century] and /*mengwa*/<sup>12</sup> “lack”) like Ibero-Romance. The outcome of JANUA(M) is surprising, however, since the development of both of these words (11) is quite different.

- (11) a. JANUA(M) > \**dʒanwa* > PGc \**dʒana* > OGC *dʒã*  
 b. MINUARE > \**menwar* > PGc \**mingwar* > OGC *mingar* (not \**mëar*, like *dʒëer* in (12a))

<sup>11</sup> Spanish /*mengwar*/ was remade on either the noun /*mengwa*/ “lack” or stressed verb forms such as /*mengwa*/ “he lessens,” neither subject to raising.

<sup>12</sup> Basque has borrowed /*mengo*/ “need.”

If MINUARE (11b) is in fact native to Gascon, its development should parallel that of JANUA(M) (11a). We can explain this inconsistency by stress—that is, /w/-deletion occurs in palatal (12a) and unstressed environments (12b), and epenthesis occurs elsewhere (12c).

- (12) a. /w/-deletion (palatal environment)  $nw > n > \emptyset$ :  
 JANUARIU(M) > \*dʒanwejr(u) > \*dʒener > dʒ(è)er  
 b. /w/-deletion (unstressed syllable)  $nw > n > \emptyset$ :  
 JANUA(M) > \*dʒanwa > \*dʒana > dʒã  
 c. /g/-epenthesis (\_\_\_ [á])  $nw > \eta g(w)$ :  
 MINUARE > \*minwar > \*miŋgar > miŋgar  
 (cf. /gw/ results for OSp and OPg)

Another possibility difficult to discard is that /mingar/ was borrowed from Ibero-Romance, where  $nw > \eta gw$  is well established.

1.3.4 *\*/nw/ in Gallo-Romance: Old Occitan (Old Provençal)*. As for JANUARIU(M), the number of reflexes often cited as belonging to Old Occitan is striking. According to Wartburg (1928-), /dʒenjer/, /dʒenojer/, and the like pertain to Old Provençal, the literary language. Anglade (1921:112) cites the additional form /dʒenovjer/, for which he posits Vulgar Latin \*JANOARIUM (> /dʒenojer/, next to \*JANARIUM (> /dʒenjer/). However, it seems difficult to imagine that the glide /w/ formed from CL \*JANUARIUM would have evolved as if it were not in hiatus, lowering to /o/ in the proto-language. Instead, as the development of \*(AD)MANUIRE > amanvir, amanavir, amanoir, and so on supports, forms like /dʒen(o)vjer/ are probably the result of vowel epenthesis between original /n\_v/. The two labial segments [ov] could then contract to [o], yielding /dʒenojer/. Furthermore, it seems that all these forms derive from borrowings from Old French containing /nv/ (see the French treatment of /nw/ in later discussion). As for TENUE, by virtue of its late learned entrance into Old Occitan, \*/tenwe/ would have escaped glide deletion like that found in /dʒenjer/, undergoing metathesis of post-tonic /w/ much like learned VIDU- → vewða/vewza. Example (13) shows the outcomes of /nw/ in Gallo-Romance.

- (13) a. /w/-develarization/V-epenthesis  $nw > nv, now > no$ :  
 manwian > \*manwir (> amanvir) > \*amanowir > amanoir  
 b. One form (probably borrowed from ML) shows metathesis:  
 TENUE “thin” > tewn (cf. VIDU- > vewða)

1.3.5 *\*/nw/ in Gallo-Romance: Old French*. Standard Old French poses virtually no problems, exhibiting general  $w > v$  (14a). Old French dialects undergo vowel insertion like the Old Occitan dialects treated earlier (14b).

- (14) a. /w/-develarization *nw* > *nv*:  
 JANUARIU(M) > \*dʒenwajr > \*dʒenvjer > OFr dʒɛ̃vjer  
 b. V-epenthesis (in dialects) *nw* > *nev*:  
 Gothic manwian > \*manwir > amanevir  
 JANUARIU(M) > \*dʒenwajr > \*dʒenvjer > Northern dʒenevɛr  
 ML TENUE > \*tenwe > Northern teneve

#### 1.4 *The liquids* \*/lw/, \*/rw/

The liquids /l, r/ could combine with /w/ in CL. In all of Romance, this original /w/ developed into a labial fricative or stop in syllable onset position (e.g., SILVA(M) > \*silwa “forest” > It selva). Although not listed in Table 2 due to their few Romance correspondences, new \*/rw/ and \*/lw/ sequences developed in Proto-Romance from vowels in hiatus. Examples are limited exclusively to verbs.

- (15) PARUIT “seem” > \*parwi > OOc parék, Sd parbit, OIt parve  
 DOLUI “ail” > \*dolwi > OIt dolvi<sup>13</sup>  
 VALUI “be worth” > \*βalwi > OOc valǵí, OCat valk, OPg valve  
 VOLUI “want” > \*βolwi > OOc volk/volǵí, OCat volk

The comparative evidence in (15), for example, OIt dolvi ~ selva, suggests that ‘secondary’ \*/rw/ and \*/lw/ merged with the original sequences. Admittedly, there are very few forms to work with, but the fact that Portuguese shares the innovation *lw* > *lv* with Italian and Sardinian is probably not due to chance alone. Thus, as in (16), develarization was possibly pan-Romance. The later restriction of *rw* > *rv* to just Portuguese, Italian, and Sardinian is most likely due to the restructuring of these root-stressed roots in most of Romance on the ‘weak’ model (e.g., OSP /βalió/).

- (16) /w/-develarization *Lw* > *Lβ/Lv*:  
 VALUI “be worth” > \*βalwi > \*βalβi > OPg valve

The importance of this finding is that these verbs in Occitan and Catalan can by no means have merged with the /lv/ and /rv/ clusters (17a) and then unmerged so as to undergo the putative sound change *w* > *gw* (17b), as accounts such as Meyer-Lübke (1895) and Fouché (1931) require.

- (17) a. PARUIT > \*parwe > \*parve  
 b. \*parve > \*parwe > \*pargwe > park  
 cf. CALVA(M) > OOc/OCat kalva “bald,” not \*kalga

<sup>13</sup> /volle/, rather than the expected \*/volve/, is analogical with other strong perfects (Tekavčić 1972:265-266).



d. [VCw]:	
grimward (proper name)	OOc grimoart (proper name)
raginward (proper name)	OOc rajnoart (proper name)
manwian “prepare”	OOc amanoir/amanvir, OCat amanir
sparwareis “sparrow”	OOc esparvier, OCat esparβer
al(e)war (“all wars”)	OCat alβar (proper name)
gel(o)wira (gails wers “glad”)	OCat elβira (proper name)

Perhaps, in one last attempt to rescue the traditional account, one could somehow invoke analogy to explain the retention of /w/ in certain verb forms, and then apply Germanic  $w > gw$  here. Even if we were to accept such an imaginative scenario, it seems very unlikely that only these medial /Cw/ sequences would receive this special treatment (e.g.,  $w > gw$ ) and not  $w > v$  or  $w > o(v)$  like the rest of the forms examined earlier. Thus there are absolutely no cases of medial  $w > gw$  in Germanic loans, as seen before for the larger Latin stock.

### 3. *So where did the velar preterites come from?*

#### 3.1 *The phonetic and phonological factors: Dissimilation of \*bw > gw*

Throughout sections 1 and 2, we saw that, of all the clusters examined, the only possible candidate for any sound change  $w > gw$  in Occitan and Catalan is \*/bw/,<sup>14</sup> because all of the (admittedly very few) reflexes of this sequence contain the velar /g/. At first blush, the very circumscribed development  $bw > gw$  may raise suspicions. In this section, I demonstrate its naturalness.

Ohala (1992:340-341) argues that dissimilation is just as ‘natural’ as assimilation, occurring when the listener inappropriately corrects or cleans up the speech signal. In short, listeners always factor out normal coarticulatory effects such as palatalization or labialization, for example, the [k<sup>l</sup>] of [k<sup>l</sup>i:] “key,” perceived as /k/. If the listener corrects some intended property, however, the percept is this segment minus the property, for example, aβwi > aγwi, with ‘hypercorrection’ of the labiality of the first segment.<sup>15</sup>

In the case of outputs such as Pre-Occitan \*[aβwi] “have,” then, the listener may have misparsed the signal as /agwi/, rather than /abwi/, because of the acoustic/auditory similarity of [βw] and [γw], as seen in (19).

<sup>14</sup>Though not discussed here, the group /kw/ voices to [gw] in Western Romance.

<sup>15</sup> Alternatively, one may wish to view the change as metathesis of the labial-velar segments (i.e., [bw] or [β<sup>l</sup>]) to the more common velar-labial order (i.e., [γw] or [Y<sup>w</sup>]), cf. SANCTIFICARE > \*santiwgar > Sp santigwar “sanctify.” See Blevins and Garrett (2002) for phonetic explanations for this frequent sound change. What is important is that initial labial + labial (-velar) changes to velar + labial, with loss of the first segment’s labiality (i.e., dissimilation).

(19) speaker  $\longrightarrow$  \* $[a\beta wi]$   $\longrightarrow$  listener  
 /abwi/  $\longrightarrow$  /agwi/

Labial-labial dissimilation<sup>16</sup> is not uncommon, even within the Romance languages, as the examples in (20) demonstrate.

(20) Spanish aβwelo > aγwelo “grandfather” (some varieties)  
 Leonese \*arbolejs > \*arβweles > arγweles (place name)  
 Aragonese senebwe > seney(w)é (place name)  
 Gothic \*triwwa > triggwa “agreement”  
 Shona bwa > bya “dog”

### 3.2 The morphological factors: Analogy and extension in the perfect tense

3.2.1 *The velar preterites.* Having motivated the sound change *bw* > *gw*, we can now examine how the velar spread beyond the original labial roots. One of the verbs to undergo *bw* > *gw* was the verb HABERE, whose perfect paradigm was as in (21).

(21) HABUI \*áβwi > \*áγwi > ak  
 HABUISTI \*aβwísti > \*aγwísti > aγís(t)  
 HABUIT \*áβwet > \*áγwet > ak  
 HABUIMUS \*aβwémos > \*aγwémos > aγém  
 HABUISTES \*aβwéstes > \*aγwéstes > aγéts  
 HABUERUNT \*áβweront > \*áγweront > áγron, áγren

Dissimilation led to labial ~ velar allomorphy in verbs like HABERE, which, in addition to its possessive function, also came to be used as a perfect and future auxiliary. It is not conceptually difficult to imagine that the velar, occurring only in the original perfect tenses, came to be reanalyzed as a past-tense marker and was later extended to other ‘strong’ (i.e., root-stressed) perfects like TENERE “have, hold,” the other less frequent possessive verb, as illustrated in (22).

<sup>16</sup> A reviewer has pointed out that it may be more accurate to speak of assimilation than of dissimilation in the change *bw* > *gw*. I agree that it is hard to tease the two apart, since dissimilation always involves coarticulation (cf. the very spread-out labialization that preceded dissimilation in the well-known development of Latin QUINQUE > k<sup>w</sup>i<sup>w</sup>.ŋ<sup>w</sup>k<sup>w</sup>e > \*ki:nk<sup>w</sup>e > It tʃinkwe “five”). Since dissimilatory change in the spirit of Ohala occurs when the listener perceives this normal coarticulation as resulting from another underlying form, embracing a dissimilation account would explain why only the voiced labial [β] (and not [ð], cf. VIDUA(M) > \*βeðða > \*βeβa > Oc veva) would yield a velar when in contact with following [w]: because, given a form \*[aβwi], [aγwi], or fully assimilated \*[awwi], a speaker/listener of a language with existing /gw/ sequences could very easily infer /agwi/, which most likely would have been realized identically.

- (22) Probable analogy with high frequency verb “have”:

3PRES <u>a</u> . <u>ten</u>			
3PERF a-k		<b>ten-k</b>	
SEDI, *SEDUI “sit <sub>1PERF</sub> ”	*sew	→	OOC sek
*CREDUIT “believe <sub>3PERF</sub> ”	*krew	→	OCat krek
STETIT (modeled on DEDIT) “stand <sub>3PERF</sub> ”	estet	→	OOC esték (rare)
*QU(A)ESIT “want <sub>3PERF</sub> ”	kes (INFIN kerér)	→	OOC kerék (rare)
PARUIT “seem <sub>3PERF</sub> ”	*par(f) (INFIN parér)	→	OOC parék
VALUI “worth <sub>1PERF</sub> ”	*val(f)	→	OOC valk

Table 9: *Further extension*

Other verbs came to adopt this pattern, as shown in Table 9. Roots in /p/ like SAPUI were not subject to this extension of /g/, however, since these roots were already marked for past tense by /w/-metathesis, as in (23).

- (23) sawp      sawβem  
       sawβist    sawβetz  
       sawp      sawβron

3.2.2 *Another case of analogy with “have”: Ibero-Romance.* Recall that the high token frequency of the verb HABUI “have” was one factor that may have favored the extension of a sound pattern (here the velar) throughout other verbal paradigms. In this section, I show that careful attention to Ibero-Romance data reveals that the scope of an oft-cited sound change is narrowed when one takes into account the workings of analogy.

Many scholars try to derive all or some of the Spanish and Portuguese strong preterites by regular sound change. For instance, Fouché (1931) posits metathesis in roots ending in labial, coronal, or dorsal (i.e., velar) segments. Velar roots appear in (24).

- (24) Metathesis in preterites:  
 JACUI > \*dʒakwi > \*dʒagwi > \*dʒawgi > OSp joge, OPg dʒowge “lie<sub>1PERF</sub>”  
 PLACUI > plakwi > \*plagwi > \*plawgi > OSp ploge, OPg prowge “please<sub>1PERF</sub>”

There are obvious difficulties with this account. First, /kw/ would have had to metathesize early on in Ibero-Romance given that both Spanish and Portuguese share this development. However, metathesis must follow voicing, since /aw/ impedes voicing. Thus proto-Western-Romance would have had unmetathesized \*/dʒagwe/ (cf. OOC [dʒak] “lay”). However, there is one



problem: Ibero-Romance and possibly all of Romance seems to have merged the outcomes of Latin /kw/ and Proto-Romance (secondary) \*/kw/, e.g., COAGULARE > \*kuaglare > OSp kwaɣar, OPg kwaɣar “curdle,” \*ECCU HIC > \*aku ík > akwí > OSp/OPg akí (cf. QUEARERE > \*kwerér > OSp/OPg kerér “want”). On this account, any general /kw/ metathesis would then have had to occur in words like Sp/Pg /agwa/ (< AQUA “water”). Unfortunately, no \*/awga/ ever existed, for this form would have given OSp \*/oga/ and OPg /owga/, two completely unattested forms, as in (25).<sup>17</sup>

- (25) Problem: If JACUI > OSp dɔge, OPg dɔwge, then  
AQUA “water” > agwa > \*awga > OSp \*oga, OPg \*owga

Second, comparative evidence suggests that the putative /kw/-metathesis is not as early as its \*/pw/ counterpart, which appears in Gallo-Romance as well, as in OoC [dɔzak] “lay” next to [sawp] “knew.” How, then, do we account for the forms in (24)? Because there is no evidence for metathesis of /kw/, it seems we are dealing with analogy. In Table 10, the /o/ or /ow/ pattern, which ‘double-marked’ the past tense of labial roots like HABUI, was extended to other consonantal roots like those ending in velar in the first column.

	OSp		OPg	
HABUI “has <sub>1PERF</sub> ”	oβe		owve	
SAPUI “know <sub>1PERF</sub> ”	sope		sowbe	
PLACUI “please <sub>1PERF</sub> ”	*plage	→ ploge	*prage	→ prowge
JACUI “lie <sub>1PERF</sub> ”	*jage	→ joge	*dɔage	→ dɔwge
TRAXI “carry <sub>1PERF</sub> ”	*treɣe	→ troge	*trejɣe	→ trowge

Table 10: Analogy with early /au/ perfects, such as “have”

In Spanish, the preterite of HABUI (i.e., /oβ-/) even caused numerous roots (many very frequently used) to adopt an alternate preterite root of the shape /(V)(C)Coβ-/, as in [toβe], [soβe], and so on, as illustrated in Table 11.<sup>18</sup>

TENUI “have <sub>1PERF</sub> ”	*tene	→ toβe
*SEDUI “sit <sub>1PERF</sub> ”	*seβe	→ soβe
*CREDUI “believe <sub>1PERF</sub> ”	*kreβ	→ kroβe
STETI “stand <sub>1PERF</sub> ”	est(j)ede	→ estoβe

Table 11: Further extension in OSp

<sup>17</sup> Dialectal Pg /awga/, as the vocalism shows, is clearly late.

<sup>18</sup> It appears that the onset of the monosyllabic root was copied here, for example, (e)st-ove, t-ove, kr-ove.

Old Spanish [oβe] was the only likely model for the extensions in Table 11. The high token frequency of the “have” verb is what favored the extension of this quite particular pattern to many other verbs. The issue of whether this extension occurred all at once or rather in small steps I leave open to further investigation. In favor of the gradient approach, it is easy to conceive of proportional analogies of the sort in (26), first gaining ground with the other “have, hold” verb *tener*, just like the Occitan case.

(26) Analogy with “have” in Old Spanish:

aβer	:	tener
oβe	:	<b>toβe</b>

#### 4. *Conclusions*

Throughout this chapter, I have shown that the putative sound change  $w > gw$  in Occitan and Catalan is ruled out on several grounds. First, there are simply no instances of any general  $w > gw$  outside of these past-tense forms. Here, the only apparent examples of  $w > gw$  are of the type  $nw > ygw$ , occurring with the greatest frequency in Spanish and Portuguese. Even if one wishes to argue for general  $nw > ygw$  in Occitan, it remains a mystery why the sound change only occurs in one form (i.e., *mingar*) in Gascon, and is apparently limited to stressed syllables before /a/. Because all verbs forms in the perfect system lack this context, however, this observation brings us no closer to understanding the velar preterites.

Proponents of the traditional account have often pointed out the similarity between the word-initial ‘fortition’ of /w/ in Germanic loanwords (e.g., *gwerra*) and the velar preterites. However, our fairly exhaustive examination of medial /Cw/ sequences (in both native words and later borrowings) has revealed no  $w > gw$  in word-medial position even here. On the alternative account, however, only labial stems like DEBUI “owe” and HABUI “have” dissimilate  $bw > gw$  and later extend this new velar past-tense marker to other verbs. The high token frequency of the verb HABUI “have” is one factor that favored the extension of the velar throughout other verbal paradigms.

As historical linguists in search of clean solutions to often dirty problems, we may be seduced by the first elegant analysis that presents itself, and cling to it for years until new data or better frameworks emerge. In this chapter, a comparative study of Ibero-Romance not only brings to light the probable influence of “have” within the verbal paradigm, but, even more importantly, also demonstrates the effect analogy can have on the formulation of sound changes.

## REFERENCES

- Anglade, Joseph. 1921. *Grammaire de l'ancien provençal ou ancienne langue d'oc*. Paris: Librairie C. Klincksieck.
- Badía Margarit, Antonio. 1951. *Gramàtica històrica catalana*. Barcelona: Noguer.
- Blevins, Juliette & Andrew Garrett. 2002. "The Evolution of Metathesis." Ms., University of California, Berkeley.
- Corominas, Joan. 1980-1991. *Diccionari etimològic i complementari de la llengua catalana*, vol. I. Barcelona: Curial/La Caixa.
- De B. Moll, Francesc. [1952] 1991. *Gramàtica històrica catalana*. Reprint, Valencia: Universitat de València.
- Devine, A. M. & Lawrence D. Stevens. 1977. *Two Studies in Latin Phonology*. Saratoga: Anma Libri.
- Fouché, Pierre. 1931. *Le verbe français: étude morphologique*. Paris: Faculté des Lettres de l'Université de Strasbourg.
- Klausenburger, Jürgen. 1984. "The Morphology of the Velar Insert in Romance Verbs". *Romanitas: Studies in Romance linguistics* ed. by Ernst Pulgram, 132-151. Ann Arbor: University of Michigan.
- Lausberg, Heinrich. 1966. *Lingüística románica. Segunda parte: Morfología*. Madrid: Gredos.
- Meyer-Lübke, W. 1890. *Grammaire des langues romanes*, vol. I, *Phonétique*. Paris: Welther.
- , 1895. *Grammaire de langues romanes*, vol. II, *Morphologie*. Paris: Welther.
- Ohala, John. 1979. "The Contribution of Acoustic Phonetics to Phonology". *Frontiers of Speech Communication Research* ed. by B. Lindblom & S. Öhman, 355-363. New York: Academic Press.
- , 1992. "What's Cognitive, What's Not, in Sound Change". *Lingua e Stile* 27:3.321-361.
- Paden, William D. 1998. *An Introduction to Old Occitan*. New York: Modern Language Association of America.
- Tekavčić, Pavao. 1972. *Grammatica storica dell'italiano. Volume 1: Fonematica*. Bologna: Mulino.
- Wartburg, Walter von. 1928-. *Französisches etymologisches Wörterbuch*. Bonn: Klopp.
- Williams, Edwin. 1938. *From Latin to Portuguese*. Philadelphia: University of Pennsylvania Press.

# THE PHONOLOGICAL ROLE OF PARADIGMS THE CASE OF INSULAR CATALAN\*

MARIA-ROSA LLORET  
*Universitat de Barcelona*

## 0. *Introduction*

Over the years, many cases have been brought to light in which the regular phonology misapplies in certain morphological environments. This issue was dealt with by means of morphologically conditioned rules in standard generative phonology, while later approaches made use of cyclic organization. Yet another approach has been to resort to structural differences encoded in the representations, such as the display of phonologically null elements with special concomitant phonological effects. The formal problems associated with these approaches have been discussed at length in the literature, but this kind of morphology-phonology interaction remains a challenge for alternative views.

In the parallel version of Optimality Theory, the notion of output-output correspondences is applicable to these cases. Within this view, there have been different proposals. The first one goes back to the notion of ‘metrical consistency’ (“Every morpheme must be as metrically consistent as possible,” Burzio 1994:228). Burzio’s (e.g., 1994) work and Kenstowicz’s (1996, 2002) work on paradigmatic uniformity and contrast share the idea that morphologically related forms create a network of possible phonological influences, in symmetric relation. Another approach is Benua’s (1997) Transderivational Correspondence Theory, which is an asymmetric model. In

---

\* This work has been supported by the Departament d’Universitats, Recerca i Societat de la Informació, Generalitat de Catalunya (Research Group 2001SGR0004), and by the Spanish Ministerio de Ciencia y Tecnología (BFF2001-3798). Previous insights on this topic were presented at the Troisième Forum International de Morphologie (Villeneuve d’Ascq, September 2002), at the Meertens Instituut (Amsterdam, March 2003), and at the International conference From Representations to Constraints (Toulouse-Le Mirail, July 2003). I am grateful to all these audiences, as well as to the audience at LSRL 33 and to the editors and the anonymous reviewers for this volume, for their insightful comments. I would also like to express my gratitude to Eulàlia Bonet, Luigi Burzio, Joan Mascaró, Clàudia Pons, and Max W. Wheeler for their suggestions, and particularly to John J. McCarthy for valuable (and generous) discussion on this material. Usual disclaimers apply.

this case, influences run in one direction only: from the morphologically simplex form to the derived complex form. More recently, McCarthy (2001) put forward a new theory of surface resemblance, the Optimal Paradigms model, which evaluates forms that are related inflectionally in symmetric relation.

The goal of this chapter is to show that the Optimal Paradigms model better captures the fact that in some languages nouns and verbs may differ phonologically in a way that is somehow connected with differences in their paradigms. I will illustrate this issue by reviewing the behavior of the inflected forms of insular Catalan with respect to vowel insertion. ‘Insular Catalan’ refers to the varieties spoken in the Sardinian town of Alguer (Italy) and in the Balearic Islands.

### 1. *The data*

In insular Catalan, that is, Alguerese (A) and Balearic (B), the first-person singular present indicative (1S.PR.IND) has no inflectional affix, like the system found in Old Catalan. Other dialects show a vocalic suffix in these forms, which is *-[u]* in Central Catalan (C) (the variety spoken in the area of Barcelona) (1a). Null affixation is also seen in regular masculine (MASC) singular nouns, which do not have any overt marker in either insular or Central Catalan (1b).

(1)	A & B	C	
a.	[mát]	[má.tu]	“kill <sub>1S.PR.IND</sub> ”
	[pás]	[pá.su]	“pass <sub>1S.PR.IND</sub> ”
	[kánt]	[kán.tu]	“sing <sub>1S.PR.IND</sub> ”
b.	[díť]	[díť]	“finger <sub>MASC</sub> ”
	[śs]	[śs]	“bone <sub>MASC</sub> ”
	[kánt]	[kán]	“song <sub>MASC</sub> ”

Insular Catalan—and Old Catalan as well—shows a rather puzzling fact for which no satisfactory explanation has been given so far: Stems with final clusters that are ill formed at the phonetic level are expected to undergo epenthesis, but 1S.PR.IND forms do not (2). In the same context, other words insert a vowel (underlined henceforth for expository reasons) (3). The examples in (2a) show sonority-increasing or plateau rhymes; the examples in (2b) show [glide + liquid] rhymes, a type of rhyme that is always rejected in Catalan, except in these verbal forms. (See more B examples in the appendix.)

(2)	A	B	C	
a.	[an.sófr]	[ən.sófr]	[ən.só.fru]	“sulfurate <sub>1S.PR.IND</sub> ”
	[é̃ntr̩]	[ə̃ntr̩]	[é̃n.tru]	“enter <sub>1S.PR.IND</sub> ”
	[ú̃mpr̩]	[ú̃mpl̩]	[óm.plu]	“fill <sub>1S.PR.IND</sub> ” <sup>1</sup>
		[pát:]	[pák.tu]	“agree <sub>1S.PR.IND</sub> ” <sup>2</sup>
b.	[res.táwr]	[rəs.táwr]	[rəs.táw.ru]	“restore <sub>1S.PR.IND</sub> ”
	[an.táwl]	[ən.táwl]	[ən.táw.lu]	“sit-down-to-table <sub>1S.PR.IND</sub> ”
(3) a.	[só.fra]	[só.frə]	[só.frə]	“sulfur <sub>MASC</sub> ”
	[sén.trá]	[sén.trə]	[sén.trə]	“center <sub>MASC</sub> ”
	[ám.pra]	[ám.plə]	[ám.plə]	“wide <sub>MASC</sub> ”
b.	[líw.rá]	[líw.rə]	[líw.rə]	“free <sub>MASC</sub> ”
	[ra.táw.lá]	[rə.táw.lə]	[rə.táw.lə]	“altarpiece <sub>MASC</sub> ”

It is worth noting that Catalan shows vowel reduction in unstressed positions. Low and mid front vowels (a, e, ε) merge as [a] in Alguerese and as schwa in Balearic and Central Catalan. Non-low back vowels also merge: in Majorca Balearic mid back vowels (o, ɔ) merge as [o], while in other dialects all non-low back vowels (o, ɔ, u) merge as [u]. These unstressed systems are responsible for setting [a] or [ə] as the default epenthetic vowel in each dialect. There are also cross-dialectal differences in the lexical distribution of stressed vowels; however, these differences are irrelevant for the aim of this chapter.

Table 1 provides the full paradigm of the present indicative tense for the sake of comparison. I have included the three conjugations, although most verbs belong to conjugation I (84%). In regular verbs, all other inflectional suffixes begin with a vowel except in athematic verbs for conjugation II, where the stem is followed by *-r/* in the infinitive, the future, and the conditional. These infinitives display regular epenthesis (e.g., /káw-r/: [káw.rá] in Alguerese, [káw.rə] in Balearic “fall<sub>INF</sub>”).<sup>3</sup> As mentioned earlier, only insular Catalan 1S.PR.IND forms do not have inflectional suffixes. In conjugations II

<sup>1</sup> In Alguerese, a historical rhotacism process applied in [obstruent + lateral] clusters; there is no actual r ~ l alternation in this context (cf. Loporcaro 1997).

<sup>2</sup> In Balearic, codas with two stops share the same point of articulation due to regressive assimilation. In Alguerese, these clusters have been historically simplified (e.g., [pát]). The same applies to laterals and nasals (e.g., *vetl* “sit-up-with<sub>1S.PR.IND</sub>” -[l:] in Balearic but -[l] in Alguerese; *condemn* “condemn<sub>1S.PR.IND</sub>” -[n:] in Balearic but -[n] in Alguerese). See Lloret (2002) for an overall description of the syllable structure in Catalan.

<sup>3</sup> In general, conjugation II is fairly problematic in Catalan, since most verbs present some kind of irregularity (cf. Wheeler 2002). Several authors suggest that, on synchronic bases, this conjugation should be withdrawn from regular paradigms (Mascaró 1986, Viaplana 1986). This interpretation is further supported by the fact that there is a clear tendency to turn conjugation II verbs to conjugation III, especially in Alguerese (e.g., *prometre* > *empromitir* “to promise,” *córrer* > *corrir* “to run”).

and III, however, there are other forms without vocalic suffixes, which may cause syllabification problems too. I will later return to this issue (nonvocalic suffixes appear in shaded cells in Table 1).

	Conjugation I			Conjugation II			Conjugation III		
	A	B	C	A	B	C	A	B	C
1S	∅	∅	[u]	∅	∅	[u]	∅	∅	[u]
2S	[as]	[əs]	[əs]	[s]	[s]	[s]	[is]	[s]	[s]
3S	[a]	[ə]	[ə]	∅	∅	∅	[i]	∅	∅
1P	[ém]	[ám]	[ém]	[ém]	[ám]/[ém]	[ém]	[ím]	[ím]	[ím]
2P	[áw]	[áw]	[éw]	[éw]	[áw]/[éw]	[éw]	[íw]	[íw]	[íw]
3P	[an]	[ən]	[ən]	[an]	[ən]	[ən]	[in]	[ən]	[ən]

Table 1: Present indicative inflectional suffixes (*S* = singular, *P* = plural)

Former analyses of the insular Catalan verbal system are based on the observation that the odd consonantal endings of 1S.PR.IND forms are possible onsets and thus their interpretation is related to this syllabic position (Mascaró 1983; Dols 1993, 2000; Dols & Wheeler 1996; Serra 1996). However, among other problems, onset analyses cannot offer a straightforward account for the overwhelming majority of coda phenomena that do take place in these verbal forms (for a review of previous analyses, see Lloret 2003).<sup>4</sup>

## 2. *The phonological evidence*

1S.PR.IND forms undergo many phonological phenomena that are associated with the coda position, summarized in Table 2. To begin with, they undergo word-final obstruent devoicing, a general phenomenon that applies to Catalan without exceptions (a).<sup>5</sup>

<sup>4</sup> For several crucial phenomena, the onset approaches mentioned here are based on incomplete data. The data in this paper are taken from different sources: the *Corpus Oral Dialectal* of the University of Barcelona (partially available on the web at: <http://www.ub.edu/linca>); Bibiloni (1983) and Pons (in progress) for Balearic Catalan; Loporcaro (1997), Bosch (2002), and Scala (2003) for Alguerese Catalan, and the literature cited before. I have also checked some of the issues in Recasens (1991) and through additional interviews to native speakers.

<sup>5</sup> Dols and Wheeler (1996) claim that in Majorca Balearic there is a systematic distinction of voicing in postconsonantal [stop + liquid] clusters. According to them, the underlying voicing of the stop is maintained in 1S.PR.IND (as in *semp* [sémbr] “SOW<sub>1S.PR.IND</sub>” vs. *empr* [émpr] “use<sub>1S.PR.IND</sub>”). This is not the usual way in which these data are reported in the literature. Recasens (1991) describes all these stops as voiceless and further mentions a low-level phonetic effect of partial devoicing of the liquid. The data of the *Corpus Oral Dialectal* also support the voiceless character of these stops.

	1S.PR.IND	Other coda contexts
a. Word-final obstruent devoicing	<i>acab</i> [p] / <i>acabar</i> [b] “I finish / to finish” <i>destorb</i> [p] / <i>destorbar</i> [b] “I hinder / to hinder” <i>cobr</i> [p] / <i>cobrar</i> [b] “I charge / to charge”	<i>tub</i> [p] / <i>tubet</i> [b] “tube / tube (diminutive)” <i>verb</i> [p] / <i>verbal</i> [b] “verb / verbal”
b. Coda v-lenition: /v/ → [w]	<i>prov</i> [w] / <i>provar</i> [v] “I prove / to prove” (Majorca B) but <i>prov</i> [f] / <i>provar</i> [v] (other insular dialects)	<i>neu(s)</i> [w] / <i>nevar</i> [v] “snow(s) / to snow” <i>viu(s)</i> [w] / <i>vivim</i> [v] “he, you live / we live”
c. /d/ → [ɾ] / V__V (A) (/d/ → [t] / __##)	<i>enfad</i> [t] / <i>enfadar</i> [ɾ] “I get angry / to get angry”	<i>fred</i> [t] / <i>freda</i> [ɾ] “cold (MASC) / cold (FEM)”
d. /l/ → [ɾ] / V__V (A)	<i>engul</i> [l] / <i>engolir</i> [ɾ] “I swallow / to swallow”	<i>sal</i> [l] / <i>saleta</i> [ɾ] “salt / salt (diminutive)”
e. Coda depalatalization: /ʎ, ɲ/ → [l, n] (A)	<i>bull</i> [l] / <i>bullir</i> [ʎ] “I boil / to boil” <i>engany</i> [n] / <i>enganyar</i> [ɲ] “I deceive / to deceive”	<i>ull(s)</i> [l] / <i>ullada</i> [ʎ] “eye(s) / look” <i>any(s)</i> [n] / <i>anyada</i> [ɲ] “year(s) / annuity”
f. Coda r-tension: /r/ → [r̥] (A, Minorca B)	<i>prepar</i> [r̥] / <i>preparar</i> [r] “I prepare / to prepare” <i>cobr</i> [r̥] / <i>cobrar</i> [r] “I charge / to charge”	<i>carta</i> [r̥] “letter” <i>per</i> [r̥] “by”

Table 2: *Coda effects*

Interestingly enough, /v/ undergoes lenition in some Majorca Balearic varieties but devoicing in all other insular varieties, as is the case in any other coda position (b). Both facts, though, are instances of coda phenomena. In addition to this, in Alguerese 1S.PR.IND forms ending in /d/ show final devoicing, as expected, although in this dialect /d/ becomes a flap between vowels, that is, in the onset position (c).<sup>6</sup> Notice that the change to a flap would be expected if the final /d/ of the verb was interpreted as the onset of an empty nucleus (in line with Dols’ 1993 and Serra’s 1996 analyses). Likewise, in Alguerese, /l/ becomes a flap between vowels, but 1S.PR.IND forms maintain the lateral in this position (d). The examples in (e) show that in Alguerese the 1S.PR.IND forms that end in /ʎ/ or /ɲ/ lose their palatal character, as they do in any other coda

<sup>6</sup> There is one single lexical exception: /r̥/ in the verbal stem of “spin” (e.g., 1S.PR.IND *fi*[r̥], 3S.PR.IND *fi*[r̥]a), but /l/ in the nominal-related words, with the expected [l] ~ [r̥] alternation (e.g., *fi*[l] “thread”, *fi*[r̥]adora “spinner”). (See also fn. 1.)



position. Finally, the examples in (f) illustrate that the insular dialects that show /r/-tension in coda position also strengthen the rhotic in 1S.PR.IND.<sup>7</sup>

1S.PR.IND forms show two further phonological peculiarities that demand an explanation. First, the 1S.PR.IND forms that end in postvocalic /r/ and /n/ maintain these consonants (4a), while in other cases these two consonants undergo deletion word finally after a stressed vowel (4b). This misapplication of the regular phonology, though, is not restricted to the first person: /r/ and /n/ are also maintained in the suffixless third persons of conjugations II and III in Balearic and Central Catalan.

(4) a.	<i>prepa</i> [r/r]	“prepare <sub>1S.PR.IND</sub> ”	<i>prepa</i> [r]es	“prepare <sub>2S.PR.IND</sub> ”
	<i>mo</i> [r/r]	“die <sub>1S.PR.IND</sub> ”	<i>mo</i> [r]im	“die <sub>1P.PR.IND</sub> ”
	<i>ma</i> [n]	“order <sub>1S.PR.IND</sub> ”	<i>ma</i> [n]es	“order <sub>2S.PR.IND</sub> ”
b.	<i>pape</i> [Ø]	“paper”	<i>pape</i> [r]era	“wastepaper”
	<i>prepara</i> [Ø]	“prepare <sub>INF</sub> ”	<i>prepara</i> [r]-ho	“prepare <sub>INF</sub> -it”
	<i>ma</i> [Ø]	“hand”	<i>ma</i> [n]ada	“handful”
c.	<i>mo</i> [r/r] (B & C)	“die <sub>3S.PR.IND</sub> ” (conjugation III) <sup>8</sup>		

Second, in Majorca Balearic there are no surface trills in the coda because /r/ is always realized as a flap in this position; however, the 1S.PR.IND forms of verbs with an underlying -/rr/ stem display final trills due to misapplication of epenthesis (5a). This exceptional phonotactic distribution of the trill—and misapplication of epenthesis—also occurs in the second and third persons of the conjugation II verb *corr-* “run” (5b), which is the only existing -/rr/ verb of conjugations II and III, which have vocalicless suffixes in the present indicative singular series.<sup>9</sup>

(5) a.	<i>ente</i> [r]	“bury <sub>1S.PR.IND</sub> ”	<i>ente</i> [r]-es	“bury <sub>2S.PR.IND</sub> ”
	<i>co</i> [r]	“run <sub>1S.PR.IND</sub> ”	<i>co</i> [r]-em	“run <sub>1P.PR.IND</sub> ”
b.	<i>co</i> [r]-s	“run <sub>2S.PR.IND</sub> ”	<i>co</i> [r]	“run <sub>3S.PR.IND</sub> ”

<sup>7</sup> The phonetic realization of /r/ in coda position presents significant cross-dialectal differences. For the purposes of this chapter, the relevant facts are that /r/ is realized as [r] in Minorca Balearic, Alguerese, and Central Catalan, while it is realized as [r̥] in the other Balearic varieties. In word-final position these realizations are stable unless they are deleted (see (4)). In word-medial position more variation is found in fast speech ([r̥], [r], and [ɹ]). Underlying /rr/ occurs phonetically as [r] and only between vowels, whether they are lexical or epenthetic, except in the verbal forms under study (see (5)).

<sup>8</sup> The verb “die” shows the innovative form [mɔ̃] for 3S.PR.IND in the city of Palma, in Majorca. In other varieties there are alternative paradigms with suffix allomorphy.

<sup>9</sup> The verb “run” shows alternative paradigms with suffix allomorphy in some varieties.

### 3. *Analysis*

The explanation I propose for the previous data is based on the following insights. First, the consonantal endings of the finite verbal forms with vocalic-less suffixes are codas because they mainly undergo typical coda phenomena. Second, the exceptional phonological behavior found in these insular verbal forms is a paradigm-uniformity effect. These two facts can be accounted for in a general view of phonology in which surface forms may violate some syllabic constraints and in which paradigms are taken into account. These are good reasons for framing the analysis within Optimality Theory, in its version of Correspondence Theory.

Before analyzing the inflected forms, we will look at a couple of cases in which paradigms do not play any role. I will use the well-known constraints stated in (6).<sup>10</sup>

- (6) a. Input-Output (IO) faithfulness constraints:
- (IO-)DEP-V: A vowel in the O corresponds to a vowel in the I.
  - (IO-)MAX-C: A consonant in the I corresponds to a consonant in the O.
- b. Constraints on syllable well-formedness:
- \*P/C: Consonants may not associate with Peak (Nucleus) nodes.
  - S(ONORITY) S(EQUENCING): Complex onsets rise in sonority, and complex codas fall in sonority.

Since in the regular phonology inputs violating SONORITY SEQUENCING trigger epenthesis, and not deletion or syllabification of C as a nucleus, the constraint ranking at work is \*P/C, SS » MAX-C » DEP-V. Two further remarks are in order here. First, consonantal cluster reduction is not driven by the sonority constraint and thus will be ignored.<sup>11</sup> Second, in Catalan, as in many other languages, the sibilant *s* has a special status in codas and does not entail epenthesis when it violates the sonority constraint. A formal solution to this well-known problem is controversial and does not shed any light on the issue here, so for the sake of expository convenience I will consider that clusters with *s* do not violate the sonority constraint. (See Bonet & Lloret 2002b for discussion on the OCP cases, which do entail epenthesis.)

Tableaux (7) and (8) depict how this ranking works in the case of non-inflected words. Tableau (7) shows that, in insular Catalan, consonantal clusters that satisfy the sonority constraint are allowed. Tableau (8) shows that

<sup>10</sup> The [glide + liquid] rhymes illustrated in (3b) would be avoided by means of a specific minimum-sonority-distance constraint. I will not examine this case further here.

<sup>11</sup> Within the framework of OT, more information on consonant deletion in Catalan can be found in Bonet and Lloret (2002a) and Pons (in progress). Partially different solutions can be found in Colina (1995), Jiménez (1999), and Dols (2000).

MAX-C is ranked above DEP-V, because the form undergoes vowel insertion when a sonority problem arises. Note that in this example Richness of the Base would provide two possible inputs: one with the final vowel and one without it. Lexicon Optimization would choose the input with the final vowel, because it gives a more harmonic mapping. But Minimal Redundancy favors the input with least underlying material, that is, the one without the final vowel, as has been assumed in (8) for the sake of illustration.

(7) *davant* “in-front”

/dəvánt/ (B)	SS	MAX-C	DEP-V
☞ a. də.vánt			
b. də.ván		*!	
c. də.ván.tə			*!

(8) *entre* “between” (Minimal Redundancy)

/əntɾ/ (Majorca B)	*P/C	SS	MAX-C	DEP-V
a. əntɾ		*!		
b. ənt.ɾ	*!			
c. ənt			*!	
☞ d. ənt.ɾə				*

In a one-by-one analysis of the words, this ranking can explain the case of epenthesis in nouns for syllabification reasons (as in *centre* /séntɾ/: [sén.tɾə], *centre-s* /séntɾ-s/: [sén.tɾəs] “center(-s)”; cf. *cèntɾ-ic* “centr-al”), but obviously it cannot also explain the failure of epenthesis in the verbal morphology of insular Catalan (as in *entr* /əntɾ/: [əntɾ] “enter<sub>1S.PR.IND</sub>” in Majorca Balearic). I am now in a position to show how underapplication of epenthesis follows from paradigm effects.

I will first consider the possibility that misapplication of epenthesis is due to paradigmatic contrast, which prohibits identical forms in a paradigm (Kenstowicz 2002). Under this view, put forward in Pons (2001) for Balearic Catalan, epenthesis could be blocked in 1S.PR.IND forms in order to avoid homophony between the first- and third-person singular of conjugation I verbs. As shown in (9), for example, in the paradigm of /kómpr/- “buy” the epenthesized first person would be homophonous with the third person, which displays the unstressed *-a* suffix.

(9)		A	B
1S.PR.IND	<i>compr</i>	[kómpr]	[kómpr]
	* <i>comprə</i>	*[kóm.pɾə]	*[kóm.pɾə]
3S.PR.IND	<i>compr-a</i>	[kóm.pɾa]	[kóm.pɾə]

There are several problems, though. To begin with, there are instances of homophony between first and third persons in tenses other than the present indicative (e.g., *compr-ava* “1S&3S.PAST,” *compr-aria* “1S&3S.CONDITIONAL,” *compr-i* “1S&3S.PR.SUBJUNCTIVE”). This problem could be resolved by appealing to the fact that in these tenses the lexicon—and not the repair strategies provided by the phonology—is responsible for the similarity; but there are other examples that cannot be handled in the same way. First, this approach cannot account for the facts of conjugation III in Alguerese (10a). Here, epenthesis in 1S.PR.IND does not apply, although paradigmatic contrast is already satisfied because conjugation III verbs display the vocalic suffix *-i* in the third person. Second, in Balearic there are verbs of conjugations II and III where epenthesis fails to apply in all singular persons that do not have vocalic suffixes (10b). Note that in this set epenthesis in 1S.PR.IND could resolve the homophony issue, but epenthesis misapplies. Epenthesis underapplication in the second persons (with alternative simplified realizations) is not related to the homophony issue either.<sup>12</sup> Third, paradigmatic contrast cannot explain the cases of nondeletion of final /r/ and /n/ in conjugation I verbs (10c). Here, underapplication of consonant deletion in 1S.PR.IND cannot be attributed to homophony, because the third person displays a vocalic suffix. The same applies to underapplication of final /r/ deletion in conjugation III verbs in Alguerese (10d). As for Balearic, the deletion of /r/ would in fact destroy the homophony between the first and third persons, but /r/ is also maintained (10e).

(10) a.	A:	“fill”	“open”	“run”
	1S.PR.IND	[úmpr]	[ópr]	[kúr]
	3S.PR.IND	[úm.pr-i]	[ó.br-i]	[kú.r-i]
	b.	B:	“open”	“run”
	1S.PR.IND	[úmpl]	[ópr]	[kór]
	2S.PR.IND	[úmpl-s]~[úm/n-s]	[ópr-s]~[ót-s]	[kór-s]
	3S.PR.IND	[úmpl]	[ópr]	[kór]
	c.	A & B:	“order”	
	1S.PR.IND	[mír/r]	[mán]	
	3S.PR.IND	[mí.r-a/ə]	[má.n-a/ə]	
	d.	A:		
	1S.PR.IND	[mór]		
	3S.PR.IND	[mó.r-i]		

<sup>12</sup> I will not discuss here the fact that the simplified forms [úms]/[úns] (with nasal place assimilation in Majorca Balearic) and [óts] (with general stop place assimilation in insular Catalan) are preferred to the more complex [úmpls] and [óprs]. This issue is related to other coda cluster conditions, not driven by the sonority constraint, involving substantial reduction and assimilation in Balearic (see Pons in progress).

e. B:	“die”
1S.PR.IND	[mɔ̃r/r]
3S.PR.IND	[mɔ̃r/r]

The overall question that remains is why the regular phonology misapplies precisely in these verbal forms. One way of answering this question is to resort to the Optimal Paradigms (OP) model (McCarthy 2001). OP is a model that incorporates elements of Metrical Consistency (e.g., Burzio 1994), Uniform Exponence (Kenstowicz 1996), and Transderivational Correspondence Theory (Benua 1997). All these theories try to capture similarities among morphologically related words through output-output correspondences. However, as several scholars have noted, the problem is to what extent Correspondence Theory is able to impose restrictions on logically possible relations, especially within inflection, where it is not clear which form should be selected as the base for attraction. OP tries to solve this problem, and its central premises are the following (McCarthy 2001:5):

- a. Candidates consist of entire inflectional paradigms.
- b. Markedness and input-output faithfulness constraints evaluate all members of the candidate paradigm. The violation marks incurred by each paradigm member are added to those incurred by all the members.
- c. The stem (shared lexeme) in each paradigm member is in correspondence relation  $\mathfrak{R}_{OP}$  with the stem in every other paradigm member. (That is, for every candidate paradigm P there is a relation  $\mathfrak{R}_{OP}$  on  $P \times P$ .) There is no distinctive base—rather, every member of a paradigm is a base of sorts with respect to every other member.

In OP, stems standing in correspondence relation are in the output because OP establishes output-output correspondences. Thus, whether the input stem loses or adds a segment in the phonetic form, the part of the surface inflected form that precedes the inflectional suffixes is identified as the base of paradigmatic relations. A similar distinction between input stems and prosodized output stems is proposed in Itô and Mester (1997) for composition and in Downing (1999) for truncation.

The OP model presupposes that nominal and verbal morphology may play a different role in determining their phonological shape because they inflect differently (McCarthy 2001:11). In languages like Catalan, where nominal inflection (with a maximum of four inflected forms) is quite limited when compared to verbal inflection (with 44 inflected forms), this thesis suggests

that paradigms have the potential to explain the phonological differences between nouns and verbs. This is the line of research that I will pursue next.

The OP approach adds a new type of constraints, those that govern the correspondence relation between the output stems of the inflected forms of a paradigm. The differences between output stems regarding the presence or absence of segments are governed by OP-MAX, which penalizes members of a paradigm with deleted segments, and OP-DEP, which penalizes members with inserted segments. For the purposes of this chapter, the relevant constraint is OP-DEP-V, which controls alternations within the paradigm with respect to vowel insertion (the output stem in each paradigm member is in correspondence relation  $\mathfrak{R}_{OP}$  with the output stem in every other paradigm member with respect to DEP-V). In insular Catalan, the ranking of OP-DEP-V above SONORITY SEQUENCING is responsible for the blocking of epenthesis in 1S.PR.IND. Tableau (11) illustrates this point (for expository reasons, I only evaluate the present indicative; the evaluation of the full verbal paradigm would not alter the results). In (11), OP-DEP-V overrides the imperatives of the sonority constraint. It penalizes five times candidate (11c), with epenthesis, because the first member of this candidate paradigm (i.e., 1S.PR.IND) contains a final vowel in the stem, while the other five members do not.<sup>13</sup> Candidate (11d), which satisfies the sonority constraint as well as OP-DEP-V by optimizing epenthesis, is discarded because of the highly ranked markedness constraint \*AA, which militates against certain hiatus. Although the syllabification of adjacent vowels is a complex issue in Catalan, for the purpose of this chapter I consider that \*AA prohibits unstressed [ə.ə]/[a.a] sequences, which is categorical in insular Catalan. That is, in the paradigm candidate (11d), the members with [ə.ə] sequences violate \*AA but the ones with [ə.á] sequences do not. Significantly, this provides evidence for a prediction made by OP, namely, that there are no true cases of underapplication; underapplication is only possible when overapplication is blocked by a highly ranked markedness constraint (in the following tableaux ‘]’ marks the right margin of the output stems standing in correspondence).

<sup>13</sup> OP-DEP-V scores one violation for each pair of forms within a paradigm and the correspondence relation is fully symmetric. Thus, in (11c) there is one violation for each of the five [án.trə]  $\mathfrak{R}_{OP}$  [ántr] relations where the final vowel of [án.trə] does not have a correspondent in [ántr]. Its symmetric counterpart, [ántr]  $\mathfrak{R}_{OP}$  [án.trə], incurs five OP-MAX-V violations, which are not considered here.

## (11) Basic ranking: \*P/C, \*AA, OP-DEP-V » SS » DEP-V

/óntr/ “enter” (conjugation I, Majorca B)	*P/C	*AA	OP-DEP-V	SS	DEP-V
☞ a. <óntr], ón.tr]əs, ón.tr]ə, ən.tr]ám, ən.tr]áw, ón.tr]ən>				*	
b. <ón.tr], ón.tr]əs, ón.tr]ə, ən.tr]ám, ən.tr]áw, ón.tr]ən>	*!				
c. <ón.trə], ón.tr]əs, ón.tr]ə, ən.tr]ám, ən.tr]áw, ón.tr]ən>			5*! (1x5)		*
d. <ón.trə], ón.trə].əs, ón.trə].ə, ən.trə].ám, ən.trə].áw, ón.trə].ən>		*,* *!			*,*,* *,*,*

Nouns, with a paradigm of two inflectional forms (<singular, plural>), undergo epenthesis because epenthesis levels the paradigms in the other direction (12). Here, candidate (12c), with epenthesis in both forms, satisfies OP-DEP-V because all members of the paradigm contain a vowel, where it is needed to satisfy the sonority constraint. Candidate (12c) wins although it violates twice the IO constraint DEP-V.

## (12)

/séntɾ/ “center <sub>MASC</sub> ” (Balearic)	*P/C	OP-DEP-V	SS	DEP-V
a. <séntɾ], séntɾ]s>			*,*!	
b. <сэн.тɾ], сэн.тɾ]s>	*,*!			
☞ c. <сэн.тɾə], сэн.тɾə]s>				*,*

The nominal paradigms of adjectives create further complications. Adjectives have a maximum of four inflected forms (<masculine singular, masculine plural, feminine singular, feminine plural>). The regular feminine suffix is *-a* ([ə] in Balearic, [a] in Algerese). The full paradigm of an adjective like /áspr/ “rough” contains two masculine forms that cause syllabification problems (/áspr/, /áspr-s/) and two feminine forms without syllabification problems (/áspr-ə/, /áspr-ə-s/ in Balearic; /áspr-a/, /áspr-a-s/ in Algerese). At this point, the analysis wrongly chooses candidate (13a) as the winner, instead of the grammatical candidate (13d), with epenthesis in the masculine forms (the grammatical candidate is indicated with the symbol ‘☞’ in the tableau).

## (13)

/áspr/ “rough” (Balearic)	*P/C	*AA	OP-DEP-V	SS	DEP-V
☞ a. <áspr], áspr]s, ás.pr]ə, ás.pr]əs>				*,*	
b. <ás.pɾ], ás.pɾ]s, ás.pɾ]ə, ás.pɾ]əs>	*,*!				
c. <ás.pɾə], ás.pɾə]s, ás.pɾə]ə, ás.pɾə]əs>		*,*!			4*
☞ d. <ás.pɾə], ás.pɾə]s, ás.pɾ]ə, ás.pɾ]əs>			4*! (2x2)		*,*

One possible explanation is to relax the family of OP constraints by acknowledging specific subsets of the paradigms, for example, <masculine singular, masculine plural> and <feminine singular, feminine plural>. This is in fact the solution put forward in Bonet, Lloret, and Mascaró (2003). Although this would act as a solution and would not alter the results in verbal morphology, it needs sufficient independent empirical support. A more challenging proposal is to relate singular and plural forms with an output-output ‘asymmetric’ correspondence relation, BASE-Identity(singular→plural) (after Kenstowicz 1996, Benua 1997), based on the fact that plurals—but not other inflected forms—are formed over freestanding output forms, that is, the singular words. However, the BASE must also contain a subset of the grammatical features of the derived form (Kager 1999:282) and, according to the traditional view, singular and plural forms are not compositionally related because of a conflict of inflectional features ([–plural] vs. [+plural]). Presumably, though, it is also possible to analyze singular forms (which never show overt inflectional markers) as being not marked for the number category; adopting this approach, there is a single feature for number, that is, [plural]. The main issue here is morphological, and for this reason I will not discuss it further.

The role of OP is to homogenize the output members of a paradigm, but it was previously shown that several coda phenomena that apply to 1S.PR.IND destroy complete uniformity. The OP model predicts that the high ranking of certain markedness constraints should ensure that OP faithfulness is not always perfect. I will illustrate this point next with the case of final devoicing. The analysis of final devoicing in terms of positional markedness has been developed in the literature according to the ranking in (14) (after Itô & Mester 1998; see also Kager 1999).<sup>14</sup>

- (14) Final devoicing: \*VOICEDCODA » IDENT(voice) » \*VOICEDOBSTRUENT
- \*VOICEDCODA (\*VCDCODA): Coda obstruents are voiceless.
  - (IO-)IDENT(voice): The specification for voice of an I must be preserved in its O correspondent.
  - \*VOICEDOBSTRUENT (\*VCDOB): Voiced obstruents are prohibited.

---

<sup>14</sup> Despite the too-many-solutions problem entailed by the positional markedness analysis of final devoicing, I follow this view based on the observations made by different scholars about existing changes in strong positions (e.g., onsets), which are not expected to occur under positional faithfulness (Zoll 1998; Steriade 2001). An alternative positional faithfulness analysis in line with Lombardi’s (2001) work, though, is possible here and will not alter the results. (See also fn. 17.)



The tableau in (15) shows this ranking at work with the OP constraints in the case of nouns. For the purposes of this chapter, the ranking of the IO constraint DEP-V above the IO constraint IDENT(voice) discards candidate (15b), with a possible epenthesis to satisfy OP-IDENT(voice), which controls alternations within the paradigm with respect to voicing.<sup>15</sup> Candidates (15a,c,d) are presented with additional stop place assimilation (resulting in an affricate), which is categorical in insular Catalan. The OP-IDENT(voice) constraint enforces the same voice feature in all paradigm members, but its effects are not visible due to its low ranking (see (16)).<sup>16</sup>

(15)

/túb/ “tube <sub>MASC</sub> ”	*VCD CODA	OP- DEP-V	SS	DEP-V	ID(vc)	OP- ID(vc)	*VCD OB
a. <túb], túd]s>	*,*!						*,*
b. <tú.bə], tú.bə]s>				*,*!			*,*
c. <túp], táf]s>					*,*		
d. <túb], táf]s>	*!				*	2* (1x1x2)	*

In the verb, the high ranking of the markedness constraint \*VOICEDCODA ensures final devoicing in 1S.PR.IND (16). This is so because OP-IDENT(voice) is ranked low, at least, below the IO constraint IDENT(voice). This can be seen by comparing the evaluations of candidates (16d) and (16e), which fare well even with respect to the sonority constraint. The crucial domination of OP constraints reinforces the idea that “OP faithfulness constraints are true OT constraints, in the sense that they are ranked within a hierarchy and are violable under crucial domination” (McCarthy 2001:32).

<sup>15</sup> Although for the sake of this chapter I assume that the ranking of IO-DEP-V above IO-IDENT(voice) penalizes epenthesis as a repair strategy to maintain voicing, in a more thorough analysis FINAL-C can do the same job. In this case, IO-DEP-V could be ranked lower (cf. Bonet & Lloret 2002a).

<sup>16</sup> In (15d) there are two violations of OP-IDENT(voice): one for the [túb]  $\mathfrak{R}_{OP}$  [táf] relation with respect to [-voice] and another one for its symmetric counterpart, [táf]  $\mathfrak{R}_{OP}$  [túb], with respect to [+voice].

(16)

/kóbr/ “charge” (Majorca B, conj. I)	*VCD CODA	*AA	OP- DEP-V	SS	DEP-V	ID(vc)	OP- ID(vc)	*VCD OB
a. <kóbr], kó.br]əs, kó.br]ə, ko.br]ám, ko.br]áw, kó.br]ən>	*!			*				*,*,* *,*,* *,*,*
b. <kó.brə], kó.br]əs, kó.br]ə, ko.br]ám, ko.br]áw, kó.br]ən>			5*! (1x5)		*			*,*,* *,*,* *,*,*
c. <kó.brə], kó.brə].əs, kó.brə].ə, ko.brə].ám, ko.brə].áw, kó.brə].ən>		*,* *!			*,*,* *,*,*			*,*,* *,*,* *,*,*
d. <kópr], kó.br]əs, kó.br]ə, ko.br]ám, ko.br]áw, kó.br]ən>				*		*	10* (1x5x2)	*,* *,*,* *,*,*
e. <kópr], kó.pr]əs, kó.pr]ə, ko.pr]ám, ko.pr]áw, kó.pr]ən>				*		*,*,* *,*,*!		

Similar analyses can be provided for other coda effects, such as depalatalization in Alguerese (17a), r-tension in some insular dialects (17b), and /d/ and /l/ not turning to [r] in Alguerese (17c). The examples in (17) are from Alguerese.

- (17) a. /búλ/ “boil”: <búλ], búλ]is, búλ]i ...>
- b. /mír/ “look at”: <mír], mír]as, mír]a ...>
- c. /púð/ “stink”: <púð], púr]is, púr]i ...>
- /angúλ/ “swallow”: <aŋgúλ], aŋgúr]is, aŋgúr]i ...>

Paradigmatic effects also play a crucial role in the exceptional behavior of the verbal forms that was previously mentioned with respect to word-final /n/ and /r/ (see (4)), and the distribution of the trill (see (5)). In general, consonant deletion and the distribution of rhotics entail further complications in the system of insular Catalan, and for this reason I leave these issues open to further research.

4. *OP and dialectal variation*

The analysis in this section illustrates a case of dialectal variation due to differences in the ranking of OP constraints. The example comes from stems ending in /v/. In Catalan, postvocalic voiced labial fricatives weaken in coda position (18a). In nominal inflection this change takes place straightforwardly

(18b). However, in insular Catalan, verbs with a final /v/ stem show 1S.PR.IND forms either with the glide or with devoicing (18c). This is not free variation but is conditioned geographically and across time. Older people in Majorca show lenition; other insular varieties show final devoicing.<sup>17</sup>

- (18) a. Regular phonology /v/: [w] in postvocalic coda (A & B & C)  
 /név/: [néw] “snow”  
 [né.v-i] “snow<sub>3S.PR.SUBJUNCTIVE</sub>”
- b. Nominal morphology /v/: [w] in postvocalic coda (A & B & C)  
 /kráv/: [kráw], [kráw-s] (A) “nail(-s)”  
 /kláv/: [kláw], [kláw-s] (B & C) “nail(-s)”
- c. 1S.PR.IND in postvocalic coda: variation (A & B)
- /v/: [w] (old Majorca B)  
 /kláv/: [kláw] “knock-in<sub>1S.PR.IND</sub>”  
 [klá.v-əs] “knock-in<sub>2S.PR.IND</sub>”  
 /próv/: [prów] “prove<sub>1S.PR.IND</sub>”  
 [pró.v-əs] “prove<sub>2S.PR.IND</sub>”
  - /v/: [f] (A & other B)  
 /kráv/: [kráf], [krá.v-as] (A)  
 /kláv/: [kláf], [klá.v-əs] (B)  
 /próv/: [próf], [pró.v-as] (A)/ [pró.v-əs] (B)

The weakening of the voiced labial fricative into a glide involves a minimal change: a change in obstruency (19a). The other potential candidate to undergo this minimal change is the voiced palatal fricative, which could turn into the palatal glide (19b). But this does not happen because Catalan undergoes a process of word-final tension (turning the palatal fricative into an affricate), and in insular Catalan this is a lexicalized change (19c).<sup>18</sup>

- (19) a. v → w: ±sonorant \*f → w: ±sonorant, ±voice  
 b. \*ʒ → j: ±sonorant \*ʃ → j: ±sonorant, ±voice  
 c. ʒ → dʒ → tʃ (word-final tension and final devoicing)

<sup>17</sup> These data contribute to Steriade’s (2001) perceptual P-map approach in two ways. First, P-map predicts that “modifications of voicing, especially final devoicing, should matter less than modifications of obstruency” (p. 32) because stricture differences play a major role in generating dissimilarity, but in the data under study gliding is preferred to devoicing except for paradigmatic reasons. Second, P-map assumes that innovations aim to improve a sound system in the safe regions of confusability (p. 51). The data show that speakers sacrifice devoicing by gliding, except for paradigmatic reasons. In my view, the reason is also language internal: The Catalan lexicon contains very few words with [f] as coda; thus, speakers exploit their knowledge of the system and favor the more common pattern. I nevertheless leave the issue of the specific nature of the labial fricatives open to further investigation (cf. Padgett 2002).

<sup>18</sup> In Balearic there are some limited lexical exceptions (e.g., *fuij* [fúitʃ] “flee<sub>1S.PR.IND</sub>”, but *fugir* “flee<sub>INF</sub>” and *fugitiu* “fugitive<sub>MASC</sub>” with medial [ʒ]).

It is no mere coincidence that lenition only applies to the voiced labial fricative, and the same could be said of the difference in phonological behavior between nouns and verbs. Under the analysis proposed here, lenition implies a violation of the faithfulness constraint IDENT(sonorant) and the differences between noun and verbs are due to differences in their paradigms.

I propose the ranking in (20) to account for the facts of old Majorca Balearic (i.e., the variety spoken by older people). In this variety, postvocalic *-/v/* undergoes coda lenition in both nouns and 1S.PR.IND, because the IO constraint IDENT(sonorant) is ranked below the IO constraint IDENT(voice) and the OP constraints OP-IDENT(voice) and OP-IDENT(sonorant), which control alternations within the paradigm with respect to voicing and obstruency respectively. In this variety, the aforementioned OP-IDENT constraints do not play any decisive role due to their low ranking.

- (20) Old Majorca B: /név/: [néw] “snow,” /próv/: [prów] “prove<sub>1S.PR.IND</sub>”  
 /báf/: [báf] “steam,” /rəzérv/: [rəzérf] “reserve<sub>1S.PR.IND</sub>”  
 Ranking: ID(voice) » ID(sonorant), OP-ID(voice), OP-ID(sonorant)

Tableaux (21) and (22) illustrate the case of nouns. Tableau (21) shows that the IO constraint IDENT(voice) is ranked above the IO constraint IDENT(sonorant). Candidate (21b) wins because candidates (21a,e) violate \*VOICEDCODA and (21d,f) violate IO-IDENT(voice) (once again for reasons of expository convenience I consider that the addition of *s* to a well-formed coda does not provoke a sonority problem). Candidate (21c), with epenthesis, is eliminated because it violates IO-DEP-V.

(21)

/név/ “snow”	*VCD CODA	OP- DEP-V	SS	DEP-V	ID(vc)	ID(snt)	OP- ID(vc)	OP- ID(snt)
a. <név], név]s>	*,*!							
☞ b. <név], név]s>						*,*		
c. <né.və], né.və]s>				*,*!				
d. <néf], néf]s>					*,*!			
e. <név], név]s>	*!					*		2* (1x1x2)
f. <név], néf]s>					*!	*	2* (1x1x2)	2* (1x1x2)

Note: For DEP-V » ID(vc), see (15); for OP-DEP-V » SS, see (11).

(22)

/báf/ “steam”	*VCD CODA	OP- DEP-V	SS	DEP-V	ID(vc)	ID(snt)	OP- ID(vc)	OP- ID(snt)
a. <báf], báf]s>								
b. <báw], báw]s>					*,*!	*,*		

Tableau (23) illustrates a case of postvocalic *-v/* verbs. Here, candidate (23b), with epenthesis in 1S.PR.IND, is discarded because it violates OP-DEP-V. Candidate (23d) shows one violation of the constraint \*VOICEDCODA. Candidate (23f), with overapplication of lenition for the sake of paradigmatic uniformity, is eliminated because it violates the highly ranked markedness constraint \*ONSET/w (no onsets associated to the labial glide), which is a fact for insular Catalan where there are no exceptions. At this point, candidate (23c), with overapplication of devoicing, and candidate (23a), with final devoicing in 1S.PR.IND, are discarded because they violate the IO faithfulness constraint IDENT(voice). Thus, candidate (23e), with lenition in 1S.PR.IND, wins.

(23)

/próv/ “prove” (conjugation I)	*ON/w	*VCD CODA	OP- DEP-V	SS	DEP-V	ID (vc)	ID (snt)	OP- ID(vc)	OP- ID(snt)
a. <próf], pró.v]əs, pró.v]ə, pro.v]ám, pro.v]áw, pró.v]ən>						*!		10* (1x5x2)	
b. <pró.və], pró.v]əs, pró.v]ə, pro.v]ám, pro.v]áw, pró.v]ən>			5*! (1x5)		*				
c. <próf], pró.f]əs, pró.f]ə, pro.f]ám, pro.f]áw, pró.f]ən>						*,*,* *,*,*!			
d. <próv], pró.v]əs, pró.v]ə, pro.v]ám, pro.v]áw, pró.v]ən>		*!							
e. <prów], pró.v]əs, pró.v]ə, pro.v]ám, pro.v]áw, pró.v]ən>							*		10* (1x5x2)
f. <prów], pró.w]əs, pró.w]ə, pro.w]ám, pro.w]áw, pró.w]ən>	*,* *,*,*!						*,*,* *,*,*!		

The tableau in (24) shows that verbs with postconsonantal *-v/* undergo final devoicing in 1S.PR.IND, because candidate (24d), with lenition in 1S.PR.IND, violates the sonority constraint.

(24)

/rəzérv/ “reserve” (conjugation I)	*V <sub>CD</sub> CODA	OP- DEP-V	SS	DEP -V	ID (vc)	ID (snt)	OP- ID(vc)	OP- ID(snt)
a. <rə.zérv], rə.zér.v]əs, rə.zér.v]ə, rə.zər.v]ám, rə.zər.v]áw, rə.zér.v]ən>	*!							
b. <rə.zérf], rə.zér.v]əs, rə.zér.v]ə, rə.zər.v]ám, rə.zər.v]áw, rə.zér.v]ən>					*		10* (1x5x2)	
c. <rə.zérvǝ], rə.zér.v]əs, rə.zér.v]ə, rə.zər.v]ám, rə.zər.v]áw, rə.zér.v]ən>		5*! (1x5)		*				
d. <rə.zérv], rə.zér.v]əs, rə.zér.v]ə, rə.zər.v]ám, rə.zər.v]áw, rə.zér.v]ən>			*!			*		10* (1x5x2)
e. <rə.zérf], rə.zér.f]əs, rə.zér.f]ə, rə.zər.f]ám, rə.zər.f]áw, rə.zér.f]ən>					*,*,* *,*,*!			

For the varieties in (25), with devoicing in 1S.PR.IND, I propose reranking OP-IDENT(sonorant) above the IO constraint IDENT(voice) and, crucially, above the IO constraint IDENT(sonorant). What emerges from this analysis is that OP constraints do not have to preserve the ranking of their corresponding IO faithfulness constraints.

(25) A & other B:

/név/: [néw] /báf/: [báf]  
/próv/: [próf] /rəzérv/: [rəzérf]

Ranking: OP-ID(sonorant) » ID(voice) » ID(sonorant), OP-ID(voice)

The ranking of OP-IDENT(sonorant) above the IO constraint IDENT(sonorant) does not change the results for nouns (cf. tableaux (21) and (22)) and verbs with postconsonantal -v/ (cf. tableau (24)). However, the high ranking of OP-IDENT(sonorant) forces faithful outputs through paradigms in verbs with postvocalic -v/ (26).

(26)

/prɔv/ "prove" (conjugation I)	*ON/w	*VCD CODA	OP- DEP-V	SS	DEP- V	OP- ID(snt)	ID (vc)	ID (snt)	OP- ID(vc)
a. <prɔf], prɔ.v]əs, prɔ.v]ə, pro.v]ám, pro.v]áw, prɔ.v]ən>							*		10* (1x5x2)
b. <prɔ.və], prɔ.v]əs, prɔ.v]ə, pro.v]ám, pro.v]áw, prɔ.v]ən>			5*! (1x5)		*				
c. <prɔf], prɔ.f]əs, prɔ.f]ə, pro.f]ám, pro.f]áw, prɔ.f]ən>							*,*,* *,*,*!		
d. <prɔv], prɔ.v]əs, prɔ.v]ə, pro.v]ám, pro.v]áw, prɔ.v]ən>		*!							
e. <prɔw], prɔ.v]əs, prɔ.v]ə, pro.v]ám, pro.v]áw, prɔ.v]ən>						10*! (1x5x2)		*	
f. <prɔw], prɔ.w]əs, prɔ.w]ə, pro.w]ám, pro.w]áw, prɔ.w]ən>	*,* *,*,*!							*,*,* *,*,*!	

On the whole, this dialectal variation is a typical instance of differences in the ranking of certain constraints. The fact that reranking involves an OP constraint that functions to block the wrong kind of identity within a paradigm reinforces the need to assess complete paradigms as output candidates.

## 5. Conclusion

The contrast between nouns and verbs, typical of insular Catalan as far as epenthesis is concerned, is not an odd idiosyncrasy of 1S.PR.IND, but is related to other peculiar contrasts between nouns and verbs that the language shows. The OP model succeeds in grasping these differences in a way that is somehow connected with differences in the organization of their paradigms. In addition to that, some preliminary results reported in McCarthy (2001) are fully supported by the data presented in this chapter, namely, the impossibility of true underapplication within paradigms and the possibility of OP unfaithfulness for markedness reasons. This analysis further provides a novel type of evidence for the OP model: dialectal variation due to the reranking of OP constraints.

APPENDIX

	1S.PR.IND Verbal forms	Masculine singular nominals
SAME SYLLABIFICATION	-VC <i>trob</i> [p] “find” <i>mat</i> [t] “kill” <i>neg</i> [k] “deny” <i>pos</i> [s] “put” <i>agaf</i> [f] “take” <i>estim</i> [m] “love” <i>sal</i> [l] “salt” <i>ball</i> [ʎ] “dance”	-VC <i>tub</i> [p] “tube” <i>petit</i> [t] “small” <i>sec</i> [k] “dry” <i>cos</i> [s] “body” <i>tuf</i> [f] “stink” <i>fum</i> [m] “smoke” <i>sal</i> [l] “salt” <i>ell</i> [ʎ] “he”
	-VC <sub>1</sub> C <sub>2</sub> <i>cant</i> [nt] “sing” <i>romp</i> [mp] “break” <i>enfang</i> [ŋk] “muddy” <i>allarg</i> [rk] “lengthen” <i>salt</i> [lt] “jump” <i>port</i> [rt] “bring” <i>fix</i> [ts] “fix”	-VC <sub>1</sub> C <sub>2</sub> <i>pont</i> [nt] “bridge” <i>camp</i> [mp] “field” <i>fang</i> [ŋk] “mud” <i>llarg</i> [rk] “long” <i>alt</i> [lt] “tall” <i>port</i> [rt] “harbor” <i>index</i> [ts] “index”
DIFFERENT SYLLABIFICATION	-V[r] (/r/) / -VC <sub>1</sub> C <sub>1</sub> <i>corr</i> [r] “run” <i>adopt</i> [t] “adopt” <i>inject</i> [t] “inject” <i>design</i> [n:] “design” <i>condemn</i> [n:] “condemn” <i>vetl</i> [l:] “watch over”	-V[r]ə (/r/) / -VC <sub>1</sub> C <sub>1</sub> ə <i>esquerre</i> [.rə] “left” <i>apte</i> [t.tə] “apt” <i>acte</i> [t.tə] “act” <i>signe</i> [n.nə] “sign” <i>solemne</i> [n.nə] “solemn” <i>batle</i> [l.lə] “mayor”
	-VCL <i>obr</i> [p] “open” <i>logr</i> [kr] “achieve” <i>ensofr</i> [fr] “sulfurate” <i>camufl</i> [fl] “camouflage” <i>arregl</i> [kl] “arrange”	-VC.Lə <i>pobre</i> [.brə] “poor” <i>alegre</i> [.grə] “happy” <i>sofre</i> [.frə] “sulfur” <i>rifl</i> [l.f] “rifle” <i>cicle</i> [l.kl] “cycle”
	-VC <sub>1</sub> C <sub>2</sub> L <i>entr</i> [ntr] “enter” <i>sebr</i> [mpr] “sow” <i>umpl</i> [mpl] “fill” <i>vincl</i> [ŋkl] “bend” <i>filtr</i> [ltr] “filter” <i>mostr</i> [str] “show” <i>mescl</i> [skl] “blend”	-VC <sub>1</sub> C <sub>2</sub> Lə <i>centre</i> [n.trə] “center” <i>timbre</i> [m.prə] “bell” <i>ample</i> [m.plə] “wide” <i>vincl</i> [ŋ.kl] “link” <i>filtr</i> [l.tr] “filter” <i>mestre</i> [s.tr] “teacher” <i>mascl</i> [s.kl] “male”
	-VGL <i>enlair</i> [jr] “raise” <i>lliur</i> [wr] “hand over” <i>m’entaül</i> [wl] “sit down to table”	-VG.Lə <i>aire</i> [j.r] “air” <i>lliure</i> [w.r] “free” <i>retaul</i> [w.l] “altarpiece”



## REFERENCES

- Benua, Laura. 1997. *Transderivational Identity: Phonological relations between words*. Ph.D. dissertation, University of Massachusetts, Amherst.
- Bibiloni, Gabriel. 1983. *La llengua dels mallorquins. Anàlisi sociolingüística*. Ph.D. dissertation, Universitat de Barcelona.
- Bonet, Eulàlia & Maria-Rosa Lloret. 2002a. "More on Alignment as an alternative to domains: The syllabification of Catalan clitics". (Available online at: <http://roa.rutgers.edu>, ROA No. 592-0403) To appear in *Probus* 17:1, 2005.
- , 2002b. "OCP Effects in Catalan Cliticization". *Catalan Journal of Linguistics* 1.19-39.
- Bonet, Eulàlia, Maria-Rosa Lloret, & Joan Mascaró. 2003. "Phonology-Morphology Conflicts in Gender Allomorphy". Paper presented at the 26<sup>th</sup> GLOW Colloquium, Lund, April 2003.
- Bosch, Andreu. 2002. *El català de l'Alguer*. Barcelona: Abadia de Montserrat.
- Burzio, Luigi. 1994. *Principles of English Stress*. Cambridge: Cambridge University Press.
- Colina, Sonia. 1995. *A Constraint-Based Analysis of Syllabification in Spanish, Catalan, and Galician*. Ph.D. dissertation, University of Illinois at Urbana-Champaign.
- Dols, Nicolau. 1993. *The Predictive Formalization of Consonantal Contacts in Majorcan Catalan (Empirical and theoretical bases)*. M.Phil. thesis, University of Sheffield.
- , 2000. *Teoria fonològica i sil·labificació. El cas del català de Mallorca*. Ph.D. dissertation, Universitat de les Illes Balears.
- Dols, Nicolau & Max W. Wheeler. 1996. "El consonantisme final del mallorquí i el 'llicenciament d'obertures'". *Caplletra* 19.51-63.
- Downing, Laura J. 1999. "Verbal Reduplication in Three Bantu Languages". *The Prosody-Morphology Interface* ed. by René Kager, Harry van der Hulst, & Wim Zonneveld, 62-89. Cambridge: Cambridge University Press.
- Itô, Junko & Armin Mester. 1997. "Correspondence and Compositionality: The Ga-gyô variation in Japanese phonology". *Derivations and Constraints in Phonology* ed. by Iggy Roca, 419-462. New York: Oxford University Press.
- , 1998. "Markedness and Word Structure: OCP Effects in Japanese". Ms., University of California, Santa Cruz. (Available online at: <http://roa.rutgers.edu>, ROA No. 528-0702)
- Jiménez, Jesús. 1999. *L'estructura sil·làbica del català*. Valencia & Barcelona: Institut Interuniversitari de Filologia Valenciana & Abadia de Montserrat.
- Kager, René. 1999. *Optimality Theory*. Cambridge: Cambridge University Press.
- Kenstowicz, Michael. 1996. "Base-Identity and Uniform Exponence: Alternatives to cyclicity". *Current Trends in Phonology: Models and methods* ed. by Jacques Durand & Bernard Laks, 363-393. Salford: University of Salford, European Studies Research Institute.
- , 2002. "Paradigmatic Uniformity and Contrast". *MIT Working Papers in Linguistics 42: Phonological Answers (and their Corresponding Questions)* ed. by Anikó Csirmaz, Zhiqiang Li, Andrew Nevins, Olga Vaysman, & Michael Wagner. Cambridge, MA: MIT Press.
- Lloret, Maria-Rosa. 2002. "Estructura sil·làbica". *Gramàtica del català contemporani* ed. by Joan Solà, Maria-Rosa Lloret, Joan Mascaró, & Manuel Pérez Saldanya, vol. I, 195-249. Barcelona: Empúries.

- , 2003. "On the Function of Emptiness". *Silixicales 3. Les unités morphologiques* ed. by Bernard Fradin, Georgette Dal, Françoise Kerleroux, Nabil Hathout, Marc Plénat, & Michel Roché, 108-117. Villeneuve d'Ascq: Université de Lille 3.
- Lombardi, Linda. 2001. "Why Place and Voice Are Different: Constraint-Specific Alternations in Optimality Theory". *Segmental Phonology in Optimality Theory: Constraints and representations* ed. by Linda Lombardi, 13-45. Cambridge: Cambridge University Press.
- Loporcaro, Michele. 1997. "On Vowel Epenthesis in Alguer Catalan". *Certamen Phonologicum III. Papers from the Third Cortona Phonology Meeting* ed. by Pier Marco Bertinetto, Livio Gaeta, Georgi Jetchev, & David Michaels, 205-227. Torino: Rosenberg & Sellier.
- Mascaró, Joan. 1983. "Three Spanish Variations and a Majorcan Theme". Ms., Universitat Autònoma de Barcelona.
- , 1986. *Morfologia*. Barcelona: Enciclopèdia Catalana.
- McCarthy, John J. 2001. "Optimal Paradigms" (Available online at: <http://roa.rutgers.edu>, ROA No. 485-1201). To appear in *Paradigms in Phonological Theory* ed. by Laura J. Downing, Tracy A. Hall, & Ralf Raffelsiefen. Oxford: Blackwell, 2004.
- Padgett, Jaye. 2002. "Russian Voicing Assimilation, Final Devoicing, and the Problem of [v]". Ms., University of California, Santa Cruz. (Available online at: <http://roa.rutgers.edu>, ROA No. 528-0702)
- Pons, Clàudia. 2001. "The Importance of Being Onset". *ConSole X Proceedings* ed. by Mario van Koppen, Joanna Sio, & Mark de Vos, 159-174. Leiden: Student Organization of Linguistics in Europe.
- , In progress. *Els contactes consonàntics en balear. Descripció i anàlisi*. Ph.D. dissertation, Universitat de Barcelona.
- Recasens, Daniel. 1991. *Fonètica descriptiva del català*. Barcelona: Institut d'Estudis Catalans.
- Scala, Luca, ed. 2003. *Català de l'Alguer: criteris de llengua escrita*. Barcelona: Abadia de Montserrat.
- Serra, Pep. 1996. *La fonologia prosòdica del català*. Ph.D. dissertation, Universitat de Girona.
- Steriade, Donca. 2001. "The Phonology of Perceptibility Effects: The P-map and its consequences for constraint organization". Ms., University of California, Los Angeles.
- Viaplana, Joaquim. 1986. "Morfologia flexiva i flexió verbal catalana". *Llengua & Literatura* 1.385-403.
- Wheeler, Max W. 2002. "Flexió verbal irregular i verbs defectius". *Gramàtica del català contemporani* ed. by Joan Solà, Maria-Rosa Lloret, Joan Mascaró, & Manuel Pérez Saldanya, vol. I, 647-729. Barcelona: Empúries.
- Zoll, Cheryl. 1998. "Positional Asymmetries and Licensing". Ms., Massachusetts Institute of Technology. (Available online at: <http://roa.rutgers.edu>, ROA No. 282-0998)



# A CONSTRAINT-BASED ANALYSIS OF GALICIAN *GEADA* \*

FERNANDO MARTÍNEZ-GIL

*The Ohio State University*

## 0. *Introduction*

A critical test for any particular phonological theory is its success in solving problems that defy an adequate solution in competing theories. This chapter deals with a rather basic phonological problem in the so-called ‘*geada*’ dialects of contemporary Galician, in which the velar obstruents [x] and [g] are found in (quasi-)complementary distribution in surface forms, and hence arguably must be derived from a single underlying source. Two minimal requirements for an adequate theoretical account of the *geada* facts are: (a) to identify the phonological segment underlying the two velar phones and (b) to propose a formal mechanism that achieves the mapping of this underlying unit to its surface realizations. As shown in this chapter, a standard serial account of *geada* falls short of meeting the second requirement, in that the lexical rule needed to account for the surface [x] ~ [g] distribution inevitably results in a violation of Structure Preservation, a general principle that restricts the output of lexical rules to the underlying inventory of a language. It is further shown that, although *geada* also presents considerable challenges to an analysis framed in terms of constraints, such as Optimality Theory (OT), this theoretical approach offers a clear advantage over a serial competitor in that the facts can ultimately be handled in a satisfactory way by means of local constraint conjunction.

The chapter is organized as follows. Section 1 presents the basic data on the *geada* dialects of Galician. In section 2, it is argued that a standard rule-based approach fails to provide an adequate analysis of the *geada* problem. Section 3 argues for an analysis of *geada* within the OT framework, one that proposes underlying /x/ and derives [g] after nasals within morphemes, thus accounting for the [x] ~ [g] surface distribution. My solution crucially resorts to the local conjunction of faithfulness and markedness constraints; it is argued that such a conjunction is justified. Finally, section 4 summarizes the account.

---

\* I would like to thank Sonia Colina and three anonymous reviewers for useful comments and discussion. All mistakes are my own.

1. *The basic data*

Galician, a Romance language closely related to Portuguese, is spoken by about two and a half million people in northwestern Spain. The phenomenon known as *geada*, characteristic of most western Galician dialects, came about as the result of a historical change that had taken place by the end of the 18<sup>th</sup> century, whereby the voiced velar stop /g/, realized phonetically as the spirant [ɣ] except in absolute word-initial position (i.e., after a pause) and after nasals, became the voiceless fricative [x] (Pensado 1970; Mariño Paz 1994). As a consequence, [g] and [x] (both spelled *g* in conventional writing) are found in complementary distribution in the synchronic grammar of *geada* dialects: [x] occurs after oral segments and [g] occurs after nasals. However, the synchronic prohibition against postnasal [x] is circumscribed to morpheme-internal position; in stem-initial position after a prefix-final nasal, only [x] is found. The examples in (1) through (3) belong to the southwestern, or Rías Baixas, variety (which incidentally corresponds to this author's own dialect); they are representative of all possible contexts in which [x] and [g] occur in *geada* Galician. The data in (1) illustrate the occurrence of [x] in a variety of environments: word initially before a vowels (1a), word initially before a tautosyllabic liquid (1b), word medially between vowels (1c), word medially before a tautosyllabic liquid (1d), and word medially after an oral consonant (1e).<sup>1</sup>

(1) a.	<i>gato</i>	[xáto]	“cat”	d.	<i>sigla</i>	[síxla]	“initial letter”
	<i>guerra</i>	[xé̃ra]	“war”		<i>aglomerar</i>	[axlomerár]	“to agglomerate”
	<i>gorxa</i>	[xór̃ša]	“throat”		<i>regra</i>	[rē̃xra]	“rule”
	<i>guía</i>	[xía]	“guide”		<i>milagre</i>	[miláxre]	“miracle”
	<i>gusto</i>	[xústo]	“taste”		<i>sagrado</i>	[saxrá̃do]	“holy, sacred”
b.	<i>grilo</i>	[xrílo]	“cricket”	e.	<i>erguer</i>	[erxér]	“to stand up”
	<i>gran</i>	[xrán]	“grain”		<i>xilgalo</i>	[šílxaro]	“goldfinch”
	<i>grego</i>	[xréxo]	“Greek”		<i>amargo</i>	[amárxo]	“bitter”
	<i>gloria</i>	[xlór̃ja]	“glory”		<i>rasgo</i>	[rásxo]	“feature”
	<i>glaciar</i>	[xlasjár]	“glacier”		<i>folga</i>	[fól̃xa]	“strike”
c.	<i>pega</i>	[péxa]	“magpie”		<i>trasgo</i>	[trásxo]	“elf, goblin”
	<i>meiga</i>	[méjxa]	“witch”		<i>algo</i>	[álxo]	“something”
	<i>lagoa</i>	[laxóa]	“lagoon”		<i>esganar</i>	[esxaná̃r]	“to choke with food”
	<i>fígado</i>	[fíxãdo]	“liver”				
	<i>agoiro</i>	[axó̃jro]	“omen”				

<sup>1</sup> Descriptions of the *geada* phenomenon can be found, among others, in Tuttle (1906), Schneider (1938), Zamora Vicente (1952), Carballo Calero (1968), Veiga Arias (1971, 1972), Fernández Rei (1981, 1990), Prieto (1980a, 1980b), Santamarina (1980), Schrotten (1980), and Pensado and Pensado Ruíz (1983).

As for the voiced velar stop [g], it occurs exclusively after a tautomorphemic nasal, as can be seen in the examples in (2), illustrating the occurrence of postnasal [g] in medial and final position within the root morpheme, in both underived forms (2a) and derived words (2b).

- |        |                   |             |              |
|--------|-------------------|-------------|--------------|
| (2) a. | <i>longo</i>      | [lón̄go]    | “long”       |
|        | <i>angazo</i>     | [aŋgáso]    | “rake”       |
|        | <i>domingo</i>    | [domín̄go]  | “Sunday”     |
|        | <i>lingua</i>     | [lín̄gwa]   | “language”   |
|        | <i>congrio</i>    | [kón̄gr̄jo] | “conger eel” |
|        | <i>xungla</i>     | [šún̄gla]   | “jungle”     |
| b.     | <i>mang+ueira</i> | [máŋḡejra] | “water hose” |
|        | <i>fung+ar</i>    | [fun̄gár]   | “to growl”   |
|        | <i>ping+ar</i>    | [pin̄gár]   | “to drip”    |
|        | <i>fung+oso</i>   | [fun̄góso]  | “fungous”    |
|        | <i>cong+oxa</i>   | [kon̄gós̄a] | “sadness”    |
|        | <i>ingl+és</i>    | [in̄glés]   | “English”    |

Interestingly, the restriction against the voiceless fricative [x] after nasals is not an absolute one. In fact, as shown by the data in (3), [x] occurs in such a context only when the nasal is in a different morpheme, as is evident in the prefixed words in (3), where [x] is morpheme initial, preceded by a prefix-final nasal. Thus, compare intramorphemic [g] in *ingl+és* “English” in (2b), not \*[in̄xlés], versus morpheme-initial [x] in *in+grato* [in̄xráto] “ungrateful,” not \*[in̄gráto].

- |        |                     |                 |                        |
|--------|---------------------|-----------------|------------------------|
| (3) a. | Underived word:     |                 |                        |
|        | <i>gordo</i>        | [xór̄do]        | “fat <sub>MASC</sub> ” |
|        | <i>gancho</i>       | [xáñ̄čo]        | “hook”                 |
|        | <i>grande</i>       | [xráñ̄de]       | “big, large”           |
|        | <i>gracia</i>       | [xrás̄ja]       | “charm, grace”         |
|        | <i>grato</i>        | [xráto]         | “grateful”             |
|        | <i>grávido</i>      | [xráβ̄iðo]      | “heavy”                |
| b.     | Prefixed word:      |                 |                        |
|        | <i>en+gardar</i>    | [en̄xordár]     | “to put on weight”     |
|        | <i>en+ganche</i>    | [en̄xáñ̄če]     | “hooking”              |
|        | <i>en+grandecer</i> | [en̄xráñ̄desér] | “to grow big”          |
|        | <i>con+graciar</i>  | [kon̄xrasjár]   | “to flatter, adulate”  |
|        | <i>in+grato</i>     | [in̄xráto]      | “ungrateful”           |
|        | <i>in+grávido</i>   | [in̄xráβ̄iðo]   | “weightless”           |

For completeness, the underlying inventory of standard Galician is given in (4). The units in parenthesis are generally absent in nonstandard dialects, including those exhibiting *geada*.

(4)

Obstruents:	{	p	t	č	k
		b	d	ǰ	(g)
		f	(θ), s	š	

Sonorants:	{	m	n	ñ
			l	(ʎ)
		r, r̄		

While [g] and [x] never enter into morphophonemic alternations in *geada* Galician, an intriguing property of these sounds is that they are found in complementary distribution: [g] occurs after nasals, and [x] word initially and after oral segments. The surface distribution is not perfectly complementary because [g] occurs after nasals in the same morpheme, whereas [x] occurs after nasals in a different morpheme.

Although the *geada* facts have been studied from a variety of perspectives, to date, no formal analysis is available in the literature.<sup>2</sup> Yet, the data in (1) through (3) are interesting from a theoretical point of view, because what in principle might resemble a relatively innocuous phonological problem turns out to present formidable challenges to formal description, whether rule based or constraint based. A fundamental claim of this chapter is that surface [g] and [x] in *geada* varieties are derived from a single underlying velar consonant, essentially corresponding to the standard Galician voiced velar stop enclosed in parentheses in (4). My main goals are: (a) to determine the phonological nature of this underlying segment and (b) to propose the formal mechanisms that account for its surface realization in *geada* Galician.

## 2. A rule-based analysis

Let us consider, first, a rule-based approach to *geada*. In principle, the data in (1) through (3) are susceptible to an analysis in terms of three logically possible hypotheses, summarized in (5).<sup>3</sup>

<sup>2</sup> From the point of view of phonotactics, *geada* Galician differs significantly from Spanish in that it allows onset [x]-liquid clusters (cf. (1b) and (1d)), a pattern prohibited in Spanish.

<sup>3</sup> There is a fourth hypothesis that can be immediately discarded, namely, that the underlying inventory of *geada* dialects includes both the velar voiceless fricative /x/ and the voiced velar stop /g/, which happen to be found in complementary distribution within morphemes. Probably no serious phonologist would consider this alternative, for a simple reason: According to a basic methodological principle of phonological analysis, when two phonetically close sounds are in complementary distribution, and unless there is compelling evidence to the contrary, they are considered surface realizations of the same phoneme, even in the absence of morphophonemic alternations. Clearly, the burden of proof would fall on an adherent of such a hypothesis to demonstrate that [g] and [x] correspond to two distinct phonemes, and to explain their highly skewed surface distribution as something other than a peculiar accident.

(5) a. **Hypothesis A:**

The underlying inventory of the *geada* dialects contains only the voiced velar stop /g/. The voiceless fricative [x] is the allophonic realization of /g/: (a) in morpheme-initial (and, by implication, also in word-initial) position, including after heteromorphemic nasals and (b) after oral segments. Of course, such a rule would recapitulate the historical events.

b. **Hypothesis B:**

The *geada* dialects have an underlying voiceless velar fricative /x/. The voiced stop [g] is the allophonic realization of /x/ after tautomorphemic nasals. Obviously, this rule entails the historical phonemic reanalysis /g/ > /x/.

c. **Hypothesis C:**

*Geada* dialects have an underlying velar unspecified for either [voice] or [continuant], or both. A rule is then needed to generate [x] and [g] in the appropriate contexts.

If we select Hypothesis A, a rule would be required to change /g/ to [x]: (a) morpheme initially and (b) after non-nasals, as in (6). Furthermore, this analysis needs to be framed within some version of lexical phonology (Kiparsky 1982, 1985; Mohanan 1986, among others), and assign (6) to a first lexical stratum, prior to prefixation.

$$(6) /g/ \rightarrow [x] / \left\{ \begin{array}{c} \# \\ [-nasal] \end{array} \right\} \text{ ---}$$

A typical derivation for the representative items *gordo* “fat,” *longo* “long,” and *engordar* “to put on weight” would proceed as in (7). Observe that rule (6) cannot be ordered after prefixation, since then it would be bled by the latter, thus ultimately yielding the incorrect form \*[eŋgordár].

(7) URs:	[#gord+o#]	[#long+o#]	[#en [#gord+a+r#] #]
1 <sup>st</sup> stratum:			
Rule (6):	xordo	----	xordar
2 <sup>nd</sup> stratum:			
Prefixation:	----	----	enxordar
Other rules:	xórðo	lónġo	eŋxorðár
Output:	[xórðo]	[lónġo]	[eŋxorðár]

A problem for Hypothesis A is that the highly restricted surface realization of underlying /g/ is implicitly viewed as an idiosyncratic property of *geada* dialects; the possibility of relating the occurrence of the voiced velar stop to the closure and voicing of the preceding nasal is inevitably lost. Indeed, there are languages in which only stops occur in postnasal position; and in yet other languages only voiced consonants surface after nasals (cf. fn. 5). The two



sequential restrictions are apparently combined within morphemes in *geada* Galician: Only the voiced velar stop occurs after tautomorphemic nasals. For Hypothesis A such a fact would be a mere accident.

If, on the other hand, we adopt Hypothesis B, then we need a rule like (8) that changes /x/ to [g] after nasals in the same morpheme.

(8) /x/ → [g] / [+nasal] \_\_\_\_

Standard generative phonology does not offer an explicit formal mechanism that allows a phonological rule to refer exclusively to ‘morpheme-internal’ position. Such a position is targeted by rules like (8), which does not contain a morpheme boundary in its structural change. The problem here is how to prevent (8) from overapplying to inputs located across morpheme boundaries. The most straightforward way to avoid such overapplication is to place (8) in a first lexical stratum, ordered before prefixation, as in (9). Observe that if (8) were to be ordered after prefixation, it would be incorrectly fed by this morphological operation, thus yielding the wrong result \*[eŋgordár].

(9) URs:	[#xord+o#]	[#lonx+o#]	[#ɛn [#xord+a+r#] #]
1 <sup>st</sup> stratum:			
Rule (9)	-----	longo	-----
2 <sup>nd</sup> stratum:			
Prefixation:	-----	-----	enxordar
Other rules:	xórdo	lóngo	eŋxorðár
Output:	[xórdo]	[lóngo]	[eŋxorðár]

There is well-established empirical evidence in Galician that supports Hypothesis B over A. Thus, in some *geada* varieties, the tautomorphemic [-ŋg-] sequence in (2) surfaces instead as [-ŋk-]. For mnemonic convenience, let us call them the ‘K-dialects.’ In addition, there is still a small group of dialects in which [x] has been generalized to all positions, including after tautomorphemic nasals; let us label them ‘X-dialects.’ Representative examples from these two dialects are given in (10).<sup>4</sup>

(10)	K-dialects	X-dialects	
<i>longo</i>	[lónko]	[lónxo]	“long”
<i>angazo</i>	[aŋkásɔ]	[aŋxásɔ]	“rake”
<i>domingo</i>	[domínko]	[domínxo]	“Sunday”
<i>mangueira</i>	[máŋkéjra]	[máŋxéjra]	“water hose”
<i>pingar</i>	[piŋkár]	[piŋxár]	“to drip”

<sup>4</sup> For further details on the two dialects illustrated in (10), see Fernández Rei (1990) and references therein.

If we were to posit underlying /x/, surface [k] in K-dialects would be derived by a stop-formation process applying in postnasal position, surely not an unusual phenomenon across languages. If, on the other hand, we were to derive [k] from underlying /g/, it would be difficult to provide any compelling justification, whether phonetic or phonological, for a process devoicing /g/ exclusively after nasals within the same morpheme. In fact, arguments against the latter analysis can readily be found in the language. As is well known, voiced obstruents in contemporary Galician exhibit a stop-spirant surface distribution essentially analogous to that found in contemporary standard Spanish: They surface as stops after homorganic nasals and laterals; spirants, on the other hand, occur elsewhere, as illustrated in (11). Such a distribution obtains both word and morpheme internally, and across morpheme and word boundaries.

(11) Spirants:				Stops:		
<i>lobo</i>	[lóβo]	“wolf”	vs.	<i>lombo</i>	[lómbo]	“back”
<i>pode</i>	[póðe]	“(s)he can”	vs.	<i>onde</i>	[ónde]	“where”
<i>é vello</i>	[é βéjɔ]	“he is old”	vs.	<i>un vello</i>	[um βéjɔ]	“an old man”
<i>teño dor</i>	[téñɔ ðór]	“I have pain”	vs.	<i>con dor</i>	[koŋ ðór]	“with pain”

Significantly, however, nonvelar stops are never devoiced in postnasal position in K-dialects (or in any other dialects for that matter). On the other hand, a bonus of positing underlying /x/ is that we can readily characterize X-dialects as a typical case of grammar simplification: They would differ from other geada varieties to the extent that they have dropped rule (8) from the grammar.

Consider now Hypothesis C, according to which the underlying segment under investigation is unspecified for either [continuant] or [voice], or both. Let us assume, for the sake of concreteness, that the underlying source is a velar fricative unspecified for voicing. Accordingly, we need a rule such as (12) that assigns the observed surface [voice] specification in the stated environments, subject to the Elsewhere Condition (Kiparsky 1982), a general convention that imposes a disjunctive ordering on the application of two rules competing for the same target, and assigns priority to the more specific rule over the more general one (the ‘elsewhere’ case).

$$(12) \left[ \begin{array}{c} -\text{son} \\ +\text{back} \end{array} \right] \rightarrow \left\{ \begin{array}{l} \text{(a) } \left[ \begin{array}{c} +\text{voice} \\ -\text{cont} \end{array} \right] / [+nasal] \\ \text{(b) } [-\text{voice}] / \text{elsewhere} \end{array} \right\}$$

As is the case in Hypotheses A and B, the structural change (a) in (12) would have to take place in a first stratum, in order to prevent the rule from applying incorrectly when the target is located after a prefix-final nasal, a restriction that would likewise need to be imposed if the underlying segment is assumed to be unspecified either for [continuant], or for both [continuant] and [voice].

The critical issue here is that a rule-based approach that resorts to lexical phonology faces an intractable problem, regardless of which of the three hypotheses (A, B, or C) is adopted, in that all three inevitably lead to a violation of Structure Preservation, a fundamental condition on lexical operations that can be paraphrased as in (13).

- (13) Structure Preservation (Kiparsky 1982, 1985; Mohanan 1986):  
Lexical operations are 'structure preserving' in the sense that they may not introduce segments or structures that are not part of the underlying inventory of a language.

In Hypotheses A and B, Structure Preservation is disobeyed because in each case the output of lexical rules (6) and (8) would be absent from the underlying inventory of *geada* Galician. In Hypothesis C, Structure Preservation is violated regardless of whether we consider an underspecified segment as formally nondistinct from a fully specified one (as in Chomsky & Halle 1968; or Kiparsky 1982), or as formally different (Archangeli & Pulleyblank 1994; Archangeli & Langendoen 1997), because the two outputs generated by (12) are distinct, and it follows by implication that both of them cannot be underlying segments in the language.

### 3. *An OT account*

Let us explore now an analysis of *geada* within the OT framework. With respect to the nature of the underlying velar, the available options are essentially reduced to selecting one of the two velars in complementary distribution and take it as basic, as summarized earlier in (5a-b); namely, either we choose /g/, as in Hypothesis A (5a), or /x/, as in Hypothesis B (5b). As a point of departure, let us adopt Hypothesis B, and thus the working assumption that [x] is basic and [g] is derived from /x/ after tautomorphemic nasals.<sup>5</sup> The

---

<sup>5</sup> An alternative that resorts to underspecification along the lines of Hypothesis C in (5c), in which the underlying segment is a velar unspecified for either [voice] or [continuant], or both, is not a viable alternative in the OT model, because it amounts to imposing a language-particular restriction on inputs, in direct contravention of a fundamental principle of the theory, the so-called 'Richness of the Base' (cf. Prince & Smolensky 1993, Smolensky 1996, Kager 1999, Smolensky, Davidson, & Juszcyk 2000, and McCarthy 2002; for some critical comparison with other conceptions of the input, see Archangeli & Langendoen 1997). The

basic formal mechanism for deriving postnasal surface [g] from underlying /x/ rests on the interaction between the featural identity (faithfulness) constraints in (14) and the markedness constraints requiring agreement of the features [voice] and [continuant] in nasal-consonant (NC) sequences, which, for convenience, I formulate as in (15).

- (14) Input-output feature identity constraints (McCarthy & Prince 1995, 1999):
- a. Identity of [voice] (IDENT-[voice]): The specification for [voice] in an input must have a correspondent in the output.
  - b. Identity of [continuant] (IDENT-[cont]): The specification for [continuant] in an input must have a correspondent in the output.
- (15) NC markedness constraints on [voice] and [continuant] specifications:
- a. \*NC: Voiceless consonants are not allowed after nasals (Pater 1996, 1999, 2001)
  - b. \*NC<sub>[cont]</sub>: Continuant consonants are not allowed after nasals.

In fact, the constraints in (15) are particular manifestations of a more general type of constraint requiring feature agreement in consonant clusters, also known in the OT literature as ‘identical cluster constraints’ (Pulleyblank 1997) or AGREE (Lombardi 1999). We may add that other highly ranked constraints barring voiceless or continuant nasals (not discussed here) militate against satisfying the two NC markedness constraints in (15) by changing the nasal’s voicing and continuancy, thereby ensuring that whenever a violation of faithfulness to the input is needed to comply with (15), the target will be the second member of the assumed underlying /-Nx-/ sequence, not the first.<sup>6</sup> The essential idea behind (15) is that surface postnasal [g] emerges from underlying /x/ as a consequence of the assimilation of the underlying unit to the closure and voicing of the preceding nasal. Accordingly, the faithfulness constraints demanding featural identity of input and output are in conflict with the NC markedness constraints requiring that a nasal-consonant sequence agree in specification of the features [voice] and [continuant]. Now, in order to compel the emergence of [g] in postnasal position in surface forms, faithfulness to input /x/ has to be sacrificed in order to satisfy agreement in voicing and continuancy with a preceding nasal, thus suggesting the preliminary constraint

---

underspecification alternative is also problematic for another reason, namely, it immediately raises the issue of the learnability of such inputs in language acquisition.

<sup>6</sup> The prohibition against voiceless consonants following a nasal is a rather common sequential constraint across languages. It is found, for example, in Zoque (Chiapas, Mexico; Padgett 1994), Kpelle (Liberia; Padgett), (Yamato) Japanese (Itô, Mester, & Padgett 1995), Puyo Pungo Quechua (Padgett), and is characteristic of the so-called ‘nasal substitution’ phenomena in the Austronesian languages (Pater 1996, 1999, 2001, and references therein). Although less common, the prohibition against postnasal fricatives is found in Rwanda (also known as Kinyarwanda; Republic of Rwanda), and in Venda and Zulu (South Africa; see Padgett and references therein).

ranking in (16), one in which the two NC constraints dominate faithfulness to the input.

(16) Constraint ranking:  $*NC_{\sigma}$ ,  $*NC_{[cont]}$  >> IDENT-[voice], IDENT-[cont]

As illustrated in (17) for the item *longo* ‘‘long,’’ (16) correctly selects (17d) over the other three likely candidates because the latter fatally violate the higher ranked NC markedness constraints.

(17) Input: /lonxo/ ‘‘long’’

Candidates	$*NC_{\sigma}$	$*NC_{cont}$	IDENT-[voice]	IDENT-[cont]
a. lónxo	*!	*		
b. lónyo		*!	*	
c. lónko	*!			*
d. $\varnothing$ lóngo			*	*

The preceding analysis, however, presents a problem, since in Galician nonvelar obstruents are not required to agree in the features [voice] and [continuant] with a preceding nasal, whether tautomorphemic or otherwise. We thus must prevent such nonvelar postnasal obstruents from violating faithfulness to underlying form in order to comply with the NC constraints.

Surface nonvelar postnasal obstruents in Galician preserve a three-way underlying contrast: (voiceless) fricative versus voiced stop versus voiceless stop, illustrated in (18) for the two other major points of articulation, namely, coronal, as in (18a), and labial, as in (18b):<sup>7</sup>

(18) a. /-Ns-/ ~ /-Nd-/ ~ /-Nt-/: <i>can<u>s</u>o</i> [kánso] ‘‘tired’’ <i>can<u>d</u>o</i> [kándo] ‘‘when’’ <i>can<u>t</u>o</i> [kánto] ‘‘how much’’	b. /-Nf-/ ~ /-Nb-/ ~ /-Np-/: <i>en<u>f</u>ermo</i> [en <u>f</u> érmo] ‘‘sick’’ <i>lomb<u>b</u>o</i> [lómbo] ‘‘back’’ <i>camp<u>p</u>o</i> [kámpo] ‘‘field’’
--	--

As illustrated in (19), under the proposed domination of the NC markedness constraints over their IDENT counterparts, a completely faithful mapping of postnasal /s/ → [s] in an item such as *canso* ‘‘tired’’ (/kanso/ → [kánso]) in (18a) would be less optimal than the likely unfaithful /s/ → \*[d], which respects  $*NC_{\sigma}$  (19a), and /s/ → \*[t], which complies with  $*NC_{[cont]}$ , (19b), thus wrongly predicting surface neutralizations of postnasal obstruents in voicing and

<sup>7</sup> As shown earlier in (4), Galician lacks underlying voiced fricatives; on the other hand, nasal-voiced fricative clusters, such as [-Nz-] and [-Nv-], do not occur in surface forms.

continuity (here and in subsequent tableaux, ‘✕’ points at the wrong candidate; the sad face points at the intended one).

(19) a. Input: /kanso/ → \*[káŋdo]

Candidates	*NC <sub>o</sub>	*NC <sub>cont</sub>	IDENT-[voice]	IDENT-[cont]
a. ☹ kánso	*!	*		
b. ✕ káŋdo			*	*

b. Input: /kanso/ → \*[káŋto]

Candidates	*NC <sub>o</sub>	*NC <sub>cont</sub>	IDENT-[voice]	IDENT-[cont]
a. ☹ kánso	*	*!		
b. ✕ káŋto	*			*

A natural way of restricting the scope of the NC constraints to postnasal velars is to appeal to \*Dor Pl (20), a member of the family of universal markedness constraints on place of articulation that disfavors velars (cf. Prince & Smolensky 1993, and much subsequent work).

(20) \*Dor Pl: No dorsal place of articulation

A simple solution to the problem posed by the data in (18) is available in OT by resorting to the local conjunction of each of the two NC markedness constraints and \*Dor Pl, as in (21). (For convenience, we will informally refer to (21a-b) as the NC<sub>[velar]</sub> markedness constraints.)

(21) a. \*NC<sub>o</sub> & \*Dor Pl  
 b. \*NC<sub>[cont]</sub> & \*Dor Pl

The net effect of the two conjunctions in (21) is to restrict the scope of the NC constraints to postnasal velars. By general convention, a constraint conjunction is violated if and only if both constraints are violated.<sup>8</sup> Accordingly, a candidate containing the sequence [-Nx-] is ruled out by (21a), while one containing [-Nk-] is prohibited by (21b). In contrast, the sequence [-Ng-] satisfies both local constraint conjunctions, as desired. The revised partial constraint ranking is shown in (22).

<sup>8</sup> For a discussion of the justification and formal properties of constraint conjunction in OT, see, among others, Kirchner (1996), Alderete (1997), Smolensky (1997), Fukazawa and Miglio (1998), Itô and Mester (1998, 2003), Kager (1999:393-400), Fukazawa (2001), Padgett (2001), Lubowicz (2001, 2002), McCarthy (2002:17-19), and Moreton and Smolensky (2002).

(22) Constraint ranking:

\*NC<sub>g</sub> & \*Dor Pl, \*NC<sub>[cont]</sub> & \*Dor Pl >> IDENT-[voice], IDENT-[cont]

Because the scope of the NC<sub>[velar]</sub> constraints in (21) is restricted to nasal-velar consonant sequences, they do not impose a requirement that nonvelar consonants agree with a preceding nasal in voicing and continuancy. The preservation of underlying distinctions in output forms for postnasal coronals and labials illustrated in (18) simply follows from the familiar input-output faithfulness to the features [voice] and [continuant] by the optimal candidate, as illustrated in (23).

(23) a. Input: /kanso/ “tired”

Candidates	ICC-[voice] & *Dor Pl	ICC-[cont] & *Dor Pl	IDENT-[voice]	IDENT-[cont]
a. $\text{\textcircled{a}}$ kánso				
b. káŋdo			*!	*
c. káŋto				*!

Now, however, we face a difficulty of a different kind. Postnasal velar stops in Galician exhibit an underlying voicing contrast, as illustrated by the minimal pairs in (24):

(24) *manga* [máŋga] “sleeve” vs. *manca* [máŋka] “one-handed<sub>FEM</sub>”  
*bingo* [bíŋgo] “bingo” vs. *vinco* [bíŋko] “ring, hoop”  
*rango* [ráŋgo] “rank” vs. *ranco* [ráŋko] “rake”  
*flamengo* [flaméŋgo] “flamingo” vs. *flamenco* [flaméŋko] “flamenco”

The assumption of underlying /x/, in conjunction with the proposed constraints and constraint ranking (22), inevitably leads to a constraint-ranking paradox. Namely, although the analysis in (22) succeeds in mapping postnasal /x/ to [g], as can be seen in the derivation of a representative item such as *longo* in (25), it invariably results in unfaithfulness to an underlying postnasal /k/, incorrectly turning it into \*[g], thus effectively obliterating the voicing distinction illustrated in (24). The undesired result is illustrated in (26) for the item *manca* [máŋka] “one-handed<sub>FEM</sub>.”

(25) Input: /lonxo/ “long”

Candidates	*NC <sub>o</sub> & *Dor PI	*NC <sub>[cont]</sub> & *Dor PI	IDENT-[voice]	IDENT-[cont]
a. lónxo	*!	*		
b. lónɣo		*!	*	
c. lónko	*!			*
d. $\varnothing$ lóngo			*	*

(26) Input: /manka/ “one-handed<sub>FEM</sub>”

Candidates	*NC <sub>o</sub> & *Dor PI	*NC <sub>[cont]</sub> & *Dor PI	IDENT-[voice]	IDENT-[cont]
a. máŋxa	*!	*		*
b. máŋya		*!	*	*
c. $\ominus$ máŋka	*!			
d. $\times$ máŋga			*	

Observe that the alternative to inverting the ranking in (22), and adopting (27) instead, is not a viable one, for although such a move would preserve underlying postnasal /k/ in the output, as shown in (28), it wrongly favors complete faithfulness to the underlying voiceless fricative in postnasal position, as illustrated in (29). Hence the ranking paradox: No permutation of the constraints assumed so far would succeed in achieving the correct results.

(27) Alternative constraint ranking:

IDENT-[voice], IDENT-[cont] >> \*NC<sub>o</sub> & \*Dor PI, \*NC<sub>[cont]</sub> & \*Dor PI

(28) Input: /manka/ “one-handed<sub>FEM</sub>”

Candidates	IDENT-[voice]	IDENT-[cont]	*NC <sub>o</sub> & *Dor PI	*NC <sub>[cont]</sub> & *Dor PI
a. máŋxa		*!	*	*
b. máŋya	*!	*		*
c. $\varnothing$ máŋka			*	
d. máŋga	*!			



(29) Input: /lonxo/ “long”

Candidates	IDENT-[voice]	IDENT-[cont]	*NC <sub>◊</sub> & *Dor Pl	*NC <sub>[cont]</sub> & *Dor Pl
a. ✕ lónxo			*	*
b. lónɣo	*!			*
c. lónko		*!	*	
d. ☉ lónɣo	*!	*		

Suppose we change our assumption about the nature of the underlying segment, and propose instead that the voiced stop /g/ is basic, and that [x] is derived in all contexts except after tautomorphic nasals, as in Hypothesis A in (5a). A ranking in which faithfulness to input [voice] and [continuant] outranks the NC<sub>[velar]</sub> markedness constraints, as in (27), would be needed in order for underlying /g/ to survive in postnasal position, as shown in (30). An advantage of this analysis is that it also succeeds in preserving postnasal /k/, as illustrated in (31).<sup>9</sup>

(30) Input: /longo/ “long”

Candidates	IDENT-[voice]	IDENT-[cont]	*NC <sub>◊</sub> & *Dor Pl	*NC <sub>[cont]</sub> & *Dor Pl
a. lónxo	*!	*	*	*
b. lónɣo		*!		*
c. lónko	*!		*	
d. ☹ lónɣo				

(31) Input: /manka/ “one-handed<sub>FEM</sub>”

Candidates	IDENT-[voice]	IDENT-[cont]	*NC <sub>◊</sub> & *Dor Pl	*NC <sub>[cont]</sub> & *Dor Pl
a. máŋxa		*!	*	*
b. máŋɣa	*!	*		*
c. ☹ máŋka			*	
d. máŋga	*!			

<sup>9</sup> An identical result would be achieved if the ranking of the IDENT constraints and \*NC<sub>[cont]</sub> & \*Dor Pl were reversed; still, it would be imperative that \*NC<sub>◊</sub> & \*Dor Pl be ranked lower than the IDENT constraints, since otherwise we would be unable to prevent the voicing of postnasal /k/ in order for the /-Nk-/ cluster to comply with \*NC<sub>◊</sub> & \*Dor Pl.

The problem we now face is how to compel the mapping /g/ → [x] in all environments except after nasals. In order to achieve this goal, we may appeal to a markedness constraint banning surface voiced velars, stated in (32), which obviously must outrank the IDENT constraints, as in (33).<sup>10</sup>

- (32) \*Voiced velar: Voiced velars are prohibited.  
 (33) \*Voiced velar >> IDENT-[voice], IDENT-[cont] >> \*NC<sub>o</sub> & \*Dor Pl, \*NC<sub>[cont]</sub> & \*Dor Pl

This alternative fails, however, because if \*Voiced velar indeed dominates the IDENT constraints, a surface voiceless velar stop [k] will always be more faithful to input /g/ than its fricative counterpart [x] in contexts other than postnasal position, as shown in (34) for the representative item *pega* [péxa] “magpie” from (1c). (The NC<sub>[velar]</sub> markedness constraints are irrelevant here, and thus have been omitted from the evaluation.)

- (34) Input: /pega/ (→ [péxa]) “magpie”

Candidates	*Voiced velar	IDENT-[voice]	IDENT-[cont]
a. ⊖ péxa		*!	*
b. péya	*!		*
c. ✕ péka		*	
d. péga	*!		

In addition, since in this analysis \*Voiced velar outranks the IDENT constraints, it follows, by transitivity, that it must also dominate the NC<sub>[velar]</sub> markedness constraints. Now, if we were to factor \*Voiced velar into the tableau for *longo* in (30), the impossibility to obtain the correct output [loŋgo] becomes quite evident: Because outputs with either [g] or [ɣ] fatally violate \*Voiced velar, candidate (29c) \*[loŋko] would be wrongly selected as optimal.

The solution ultimately adopted here is one that posits underlying /x/, and in addition appeals to the conjunction of \*NC<sub>o</sub> and IDENT-[cont], as in (35).

- (35) \*NC<sub>o</sub> & IDENT-[cont]: An output must either agree in voicing with a preceding nasal or be faithful to the input specification for [continuant], or both.

Accordingly, the output [loŋgo] will always be a more optimal mapping of underlying /lonxo/ than its counterpart [loŋko], since the latter violates both \*NC<sub>o</sub> and IDENT-[cont], and hence also their conjunction in (35), whereas the

<sup>10</sup> Obviously, the constraint (32) would be an essential ingredient in any OT approach to the historical emergence of *geada* in Galician.

former only disobeys IDENT-[cont], and therefore satisfies (35). In short, the desired results can be obtained if (35) dominates the IDENT constraints, although (35) itself is not crucially ranked with respect to \*NC<sub>[cont]</sub> & \*Dor Pl, as suggested in (36).

(36) Constraint ranking:

\*NC<sub>[cont]</sub> & \*Dor Pl, \*NC<sub>◌</sub> & IDENT-[cont] >> IDENT-[voice], IDENT-[cont]

Essentially, the \*NC<sub>◌</sub> portion of (35) compels the mapping of /x/ to [g], but disallows [k] after nasals, as can be seen in (37), while the IDENT-[cont] portion ensures that a postnasal /k/ is faithfully preserved in surface forms, even if it disobeys the \*NC component, as shown in (38).

(37) Input: /lonxo/ “long”

Candidates	*NC <sub>[cont]</sub> & *Dor Pl	*NC <sub>◌</sub> & IDENT-[cont]	IDENT-[voice]	IDENT-[cont]
a. lónxo	*!			
b. lónyo	*!		*	
c. lónko		*!		*
d. $\varnothing$ lóngo			*	*

(38) Input: /manka/ “one-handed<sub>FEM</sub>”

Candidates	*NC <sub>[cont]</sub> & *Dor Pl	*NC <sub>◌</sub> & IDENT-[cont]	IDENT-[voice]	IDENT-[cont]
a. máñxa	*!		*	*
b. máñya	*!			*
c. $\varnothing$ máñka				
d. máñga			*!	

An immediate advantage of this solution is that the local conjunction \*NC<sub>◌</sub> & \*Dor Pl can be disposed of, since it no longer plays a critical role in the selection of the optimal output, thus resulting in a simpler analysis.<sup>11</sup> Another bonus of this alternative is that we can formally characterize the K-dialects of *geada* (in which, it should be recalled, we get [k] instead of [g] after tautomorphic nasals), by simply assuming that the constraint \*Voiced velar, which in these

<sup>11</sup> \*NC<sub>◌</sub> is freely violated in Galician in any event because the language allows homorganic nasal-voiceless consonant sequences such as [-ŋk-], [-ŋt-], and [-mp-].

varieties of course is never disobeyed, has been promoted to a higher rank than the NC<sub>[velar]</sub> markedness constraints, as illustrated in (39). Thus, the K-dialects simply differ from the standard *geada* ones in that the conjunction \*NC<sub>◊</sub> & IDENT-[cont] is no longer needed.

(39) K-Dialects: Input: /lonxo/ “long”

Candidates	*Voiced velar	*NC <sub>[cont]</sub> & *Dor Pl	IDENT-[voice]	IDENT-[cont]
a. lónxo		*!		
b. lónyo		*!	*	
c. <sup>ɸ</sup> lónjko				*
d. lóngo	*!		*	*

With respect to the solution proposed in (38), beyond the obvious success of the composite constraint \*NC<sub>◊</sub> & IDENT-[cont] in yielding the desired results, the question naturally arises of why two constraints so dissimilar in nature (i.e., \*NC<sub>◊</sub>, a markedness constraint, and IDENT-[cont], in the feature-faithfulness family) should be combined in the first place. More generally, one might ask whether local conjunction should be allowed at all. As indicated earlier (cf. fn. 7), the obvious answer to this question in OT is that constraint conjunction is justified to the extent that it provides a way to capture certain aspects of phonological structure that cannot otherwise be expressed by individual constraint interaction. In fact, constraint conjunction remains relatively controversial in OT because, on the one hand, it greatly increases OT’s descriptive power, thus predicting phonological patterns unattested in human language, and, on the other, it substantially undermines the basic OT tenet of strict constraint domination. Unfortunately, however, there is no widespread agreement at present on the nature and the type of substantive limitations that should be imposed on local conjunction. Some authors have attempted to set restrictions on the kinds constraints that may be conjoined, such as the condition that they belong to the same family (Itô & Mester 1998) or, alternatively, that they share the same domain (Fukazawa & Miglio 1998; Itô & Mester 1998, 2003), and a conjunctive analysis of faithfulness and markedness constraints along the lines of (38) is fully developed in Lubowicz (2002) in order to account for derived-environment effects in a number of languages. The only condition generally agreed on is that conjoined constraints share some common property (cf. Padgett 2001). While a detailed discussion of constraint conjunction is beyond the scope of this chapter, reasonably solid motivation can be provided in support of the language-particular combination of \*NC<sub>◊</sub> & IDENT-[cont] in the analysis of Galician *geada*. First, it meets the

requirement that the two members of a conjunction be phonetically conjoinable (McCarthy 1997). Second, the two conjoined constraints apply within morphemes, thus also satisfying the general condition that the members of a conjunction share the same domain (Smolensky 1996, 1997). Finally, and perhaps most importantly, the present analysis can be justified insofar as it succeeds in expressing two fundamental aspects of the *geada* phenomenon. First, it captures the asymmetrical distribution of velar obstruents in postnasal position: While stops exhibit a surface contrast in voicing ([-Ng-] vs. [-Nk-]), a similar contrast in continuancy is lacking (i.e., [-Ng-] vs. \*[-Ny-] or [-Nk-] vs. \*[-Nx-]). This asymmetry follows directly from the domination of \*NC<sub>[cont]</sub> & \*Dor Pl over the IDENT constraints, effectively ruling out postvelar continuants (cf. (37a-b)). And second, given the hypothesis of underlying /x/, the conjunction \*NC<sub>ç</sub> & IDENT-[cont] is fully consistent with the most basic function of constraint conjunction in OT, that of rejecting “the worse of the worse” (McCarthy 2002:18). Namely, the mapping /-Nx-/ → [-Ng-], although unfaithful, is optimal because it at least satisfies markedness, whereas its competitor /-Nx-/ → \*[-Nk-] is not because it disobeys both faithfulness and markedness.

To conclude, in contrast with the difficulties encountered with tautomorphic nasal-obstruent sequences, the task of accounting for surface [x] morpheme initially after heteromorphic nasals is straightforward, regardless of any particular assumption about the nature of the underlying segment. The reason for this is that the phonological identity of morphologically related words is handled in OT by the so-called ‘Output-to-Output Identity’ constraints, as in the theory of transderivational correspondence (Benua 2000), which establish a correspondence relation between two output forms, the base and a derived form. Such constraints militate primarily against allomorphic variation by requiring that morphologically derived words be faithful to their base. The relevant constraint is formulated in (40) (here, the base = the root morpheme).<sup>12</sup>

(40) Output-to-Output Identity (OO-IDENT):

A root in a morphologically derived word must be identical to its base.

For the purposes of this account, phonological identity between a base such as *gord-o* [xórðo] “fat<sub>MASC</sub>” and its related form *en+gord-ar* [eɲxorðár] “to gain weight” (cf. (3)) is accomplished directly by the domination of OO-IDENT over

<sup>12</sup> An essentially equivalent way of compelling identity in morphologically related words is by ‘Uniform Exponence’ (Kenstowicz 1997:39), which demands that a lexical item (stem, affix, word) have the same realization for a given property P in its various contexts of occurrence.

the NC markedness constraints, as in (41). The selection of the optimal candidate is shown in tableau (42).

- (41) OO-IDENT >> \*NC<sub>[cont]</sub> & \*Dor Pl >> IDENT-[cont] & \*NC<sub>σ</sub> >> IDENT-[voice],  
IDENT-[cont]

- (42) Input: /en+xordar/ “to gain weight” (base: [xórð-o] “fat<sub>MASC</sub>”)

Candidates	OO-IDENT	*NC <sub>[cont]</sub> & *Dor Pl	IDENT-[cont] & *NC <sub>σ</sub>	IDENT- [voice]	IDENT- [cont]
a. eɲgorðár	*! (gorð-)			*	*
b. eɲɔrðár	*! (ɔrð-)	*	*	*	
c. eɲkorðár	*! (korð-)		*		*
d. <sup>ɸ</sup> eɲxorðár		*			

As shown in (42), candidates (42a-c) fatally disobey OO-IDENT because they are unfaithful to the base [xórð-o]. Especially relevant here is that, in ruling out a form such as (45a), containing a postnasal [g], OO-IDENT avoids potential stem-initial [x] ~ [g] alternations such as [x]ord-o ~ \*en+[g]ord-ar. Observe further that a form such as *a+long-ar* [aloŋgár] “to lengthen,” assumed to derive from underlying /a+lonx-ar/ in my analysis is correctly selected over its potential competitor \*[aloŋxár] on two counts: (a) The former is faithful to the base (surface) form *longo* [lón̄go], and thus satisfies OO-IDENT, whereas the latter disobeys it; and (b) \*[aloŋxár] is less optimal than [aloŋgár] in any event for the same reason that [lón̄go] is favored over \*[lón̄xo] in (37): The choice is compelled by the proposed domination of \*NC<sub>[cont]</sub> & \*Dor Pl over the IDENT constraints.

#### 4. Conclusion

The main goal of this chapter has been to provide a formal account for the quasi-complementary surface distribution of the velars [g] and [x] found in the *geada* dialects of Galician. It was suggested, first, that the only hypothesis consistent with the data is one in which both velars are derived from a single underlying segment. A standard rule-based approach to *geada* has been shown to be untenable, because the lexical rule needed to derive the [x] ~ [g] complementary distribution inevitably leads to a violation of the principle of Structure Preservation. The *geada* data have also been shown to pose a significant challenge to a constraint-based analysis, as in the OT model. After considering each of the two allophonic realizations in (quasi-)complementary distribution as basic, it was concluded that the only viable alternative is one

that posits /x/ and derives [g] in postnasal position inside morphemes. This solution, however, is available only if we appeal to the conjunction of a markedness constraint with a faithfulness one (to \*NC<sub>o</sub> & IDENT-[cont]), a formal complexity that, I have argued, is required to match the asymmetric nature of constraints on voicing and continuancy for postnasal obstruents in *geada* Galician: Agreement in this feature applies to velars, not to other places of articulation.

#### REFERENCES

- Alderete, John. 1997. "Dissimilation as Local Conjunction". *Proceedings of NELS 27* ed. by Kiyomi Kusumoto, 17-32. Amherst, MA: GLSA.
- Archangeli, Diana & D. Terence Langendoen. 1997. "Afterword". *Optimality Theory: An overview* ed. by Diana Archangeli & Terence Langendoen, 200-215. Oxford: Blackwell.
- Archangeli, Diana & Douglas Pulleyblank. 1994. *Grounded Phonology*. Cambridge, MA: MIT Press.
- Benua, Laura. 2000. *Phonological Relations Between Words*. New York: Garland.
- Carballo Calero, Ricardo. 1968. "Geadá". *Grial* 19.99-100.
- Chomsky, Noam & Morris Halle. 1968. *The Sound Pattern of English*. New York: Harper & Row.
- Fernández Rei, Francisco. 1981. "Gheada". *Gran Enciclopedia Galega* 16.21-23.
- . 1990. *Dialectoloxía da lingua galega*. Vigo: Edicións Xerais de Galicia.
- Fukazawa, Haruka. 2001. "Local Conjunction and Extending Sympathy Theory: OCP effects in Yucatec Maya". *Segmental Phonology in Optimality Theory: Constraints and representations* ed. by Linda Lombardi, 231-260. New York: Cambridge University Press.
- Fukazawa, Haruka & Viola Miglio. 1998. "Restricting Conjunction to Constraint Families". *Proceedings of the West Coast Conference on Linguistics 9 (WECOL 96)* ed. by Vida Samiian, 102-117. Fresno, CA: Department of Linguistics, California State University at Fresno.
- Itô, Junko & Armin Mester. 1998. "Markedness and Word-Structure: OCP effects in Japanese". Ms., University of California, Santa Cruz. (Available online at: <http://roa.rutgers.edu>, ROA No. 528-0702)
- . 2003. "On the Sources of Opacity in OT". *The Syllable in Phonological Theory* ed. by Caroline Féry & Ruben van de Vijver, 271-303. New York: Cambridge University Press.
- Itô, Junko, Armin Mester, & Jaye Padgett. 1995. "Licensing and Underspecification in Optimality Theory". *Linguistic Inquiry* 26.571-613.
- Kager, René. 1999. *Optimality Theory*. New York: Cambridge University Press.
- Kenstowicz, Michael. 1997. "Base-Identity and Uniform Exponence: Alternatives to cyclicity". *Current Trends in Phonology: Models and methods* ed. by Jacques Durand & Bernard Laks, 365-395. Salford: European Studies Research Institute.
- Kiparsky, Paul. 1982. "Lexical Phonology and Morphology". *Linguistics in the Morning Calm* ed. by I. S. Yang, vol. II, 3-91. Seoul: Hanshin.
- . 1985. "Some Consequences of Lexical Phonology". *Phonology* 2.85-138.
- Kirchner, Robert. 1996. "Synchronic Chain Shifts in Optimality Theory". *Linguistic Inquiry* 27.341-350.
- Lombardi, Linda. 1999. "Positional Faithfulness and Voicing Assimilation in Optimality Theory". *Natural Language and Linguistic Theory* 17.267-302.

- Lubowicz, Anna. 2001. "Derived Environment Effects in OT". *Proceedings of the West Coast Conference on Formal Linguistics 17* ed. by Kimary N. Shahin, Susan J. Blake, & Eun-Sook Kim, 451-465. Stanford, CA.: CSLI.
- . 2002. "Derived Environment Effects in Optimality Theory". *Lingua* 112.243-280.
- Mariño Paz, Ramón. 1994. "Testemuños de gheada nos primeiros textos galegos do século XIX". *Actas do XIX Congreso Internacional de Lingüística e Filoloxía Románicas* ed. by Ramón Lorenzo, 207-222. La Coruña: Fundación Pedro Barrié de la Maza.
- McCarthy, John. 1997. "Process-Specific Constraints in Optimality Theory". *Linguistic Inquiry* 28.231-241.
- . 2002. *A Thematic Guide to Optimality Theory*. New York: Cambridge University Press.
- McCarthy, John & Alan S. Prince. 1995. "Faithfulness and Reduplicative Identity". *Papers in Optimality Theory: University of Massachusetts Occasional Papers in Linguistics 18* ed. by Jill N. Beckman, Laura Walsh Dickey, & Suzanne Urbanczyk, 249-384. Amherst: University of Massachusetts at Amherst.
- . 1999. "Faithfulness and Identity in Prosodic Morphology". *The Prosody-Morphology Interface* ed. by René Kager, Harry van der Hulst, & Wim Zonneveld, 218-309. New York: Cambridge University Press.
- Mohanan, K. P. 1986. *The Theory of Lexical Phonology*. Dordrecht: Reidel.
- Moreton, Elliot & Paul Smolensky. 2002. "Typological Consequences of Local Constraint Conjunction". Ms., Johns Hopkins University. (Available online at: <http://roa.rutgers.edu>, ROA No. 525-0602)
- Padgett, Jaye. 1994. "Stricture and Nasal Place Assimilation". *Natural Language and Linguistic Theory* 12.465-513.
- . 2001. "Constraint Conjunction Versus Grounded Constraint Subhierarchies in Optimality Theory". Ms., University of California, Santa Cruz. (Available online at: <http://roa.rutgers.edu>, ROA No. 530-0702)
- Pater, Joe. 1996. "\*NC<sub>0</sub>". *Proceedings of North East Linguistic Society 26* ed. by Kiyomi Kusumoto, 227-239. Amherst, MA: Graduate Linguistic Student Association.
- . 1999. "Austronesian Nasal Substitution and Other \*NC<sub>0</sub> Effects". *The Prosody Morphology Interface* ed. by René Kager, Harry van der Hulst, & Wim Zoenneveld, 310-343. New York: Cambridge University Press.
- . 2001. "Austronesian Nasal Substitution Revisited: What is wrong with \*NC<sub>0</sub> and what's not". *Segmental Phonology in Optimality Theory: Constraints and representations* ed. by Linda Lombardi, 59-182. New York: Cambridge University Press.
- Pensado, José L. 1970. "Interferencias estructurales castellano-gallegas: el problema de la gheada y sus causas". *Revista de Filología Española* 53.27-44.
- Pensado, José L. & Carmen Pensado Ruíz. 1983. 'Gueada' y 'geada' gallegas (*Verba* Monograph 21). Santiago de Compostela: University of Santiago de Compostela Press.
- Prieto, Domingo. 1980a. "A gheada: variantes fonéticas, distribución xeográfica e interpretación". *O Ensino* 2.39-45.
- . 1980b. "Algunhas hipóteses sobre a gheada". *Verba* 7.223-241.
- Prince, Alan & Paul Smolensky. 1993. *Optimality Theory: Constraint interaction in generative grammar* (Report No. RuCCS-TR-2). New Brunswick, NJ: Rutgers University Center for Cognitive Science. (Available online at: <http://roa.rutgers.edu>, ROA No. 537-0802)
- Pulleyblank, Douglas. 1997. "Optimality Theory and Features". *Optimality Theory: An overview* ed. by Diana Archangeli & D. Terence Langendoen, 59-101. Oxford: Blackwell.



- Santamarina, Antón. 1980. "Novas consideracións ó redor das orixes da geadá." *Verba* 7.243-249.
- Schneider, Hans. 1938. "Studien zum Galizischen des Limiabeckens (Orense-Spanien)". *Volkstum und Kultur der Romanen* 11.69-145, 193-281.
- Schroten, Jan. 1980. "Interpretación de la geadá gallega". *Verba* 7.209-222.
- Smolensky, Paul. 1996. "The Initial State and 'Richness of the Base' in Optimality Theory". (Tech. Rep. JHU-CogSci-96-4). Baltimore: Department of Cognitive Science, Johns Hopkins University. (Available online at: <http://roa.rutgers.edu>, ROA No. 154-1196)
- . 1997. "Constraint Interaction in Generative Grammar II: Local conjunction". Ms., University of Colorado.
- Smolensky, Paul, Lisa Davidson, & Peter Jusczyk. 2000. "The Initial and Final States: Theoretical implications and experimental explorations of Richness of the Base". Ms., University of Colorado. (Available online at: <http://roa.rutgers.edu>, ROA No. 428-1200)
- Tuttle, E. H. 1906. "Galician G". *Modern Philology* 4.279-280.
- Veiga Arias, Amable. 1971. "Fonología galega. Fonemática: el sistema consonántico, III". *Grial* 32.155-162.
- . 1972. "Fonología galega. Fonemática: el sistema consonántico, IV". *Grial* 37.281-291.
- Zamora Vicente, Alonso. 1952. "La frontera de la geadá". *Homenaje a Fritz Krüger, I*, 57-72. Mendoza, Argentina: Universidad Nacional de Cuyo.

# PEAK PLACEMENT IN TWO REGIONAL VARIETIES OF PERUVIAN SPANISH INTONATION\*

ERIN O'ROURKE

*University of Illinois at Urbana-Champaign*

## 0. *Introduction*

Research on Spanish intonation has shown that the alignment of peaks in a pitch contour varies according to the position, focal prominence, and even modality of an utterance. Peninsular Spanish was first described by Navarro Tomás (1944) as differentiating between two positions: The peak in nuclear (final) position in an utterance occurs within the stressed syllable, while prenuclear peaks are often realized after syllable offset in neutral declaratives. This same behavior has been observed in subsequent studies of Peninsular and Latin American Spanish, and has been shown to be characteristic of several varieties of Spanish (Hualde 2002:103).<sup>1</sup> When a particular word is emphasized under focus, the prenuclear peak is found to occur within the stressed syllable for Peninsular Spanish (Face 2002a). However, the actual placement of the peak may differ according to the Spanish variety, as seen in Sosa (1999) and McGory and Díaz-Campos (2002). In Dominican Spanish, prenuclear peaks do not shift under focus, but rather the valley is realized earlier (Willis 2003). Not only are prenuclear peaks realized earlier for declarative utterances under focus, but imperatives show an increase in the use of retracted peaks in Mexican Spanish compared to neutral declaratives (Willis

---

\* I would like to express my appreciation to Virginia Zavala Cisneros for providing me her office space at the PUCP to conduct the recordings as well as helping to recruit participants. Similarly, I would like to recognize Jean-Jeaques Decoster for granting Researcher Affiliation to the CBC and Jaime Arias Motta for referring students from ISMLAM. Thanks are also due to the participants from Lima and Cusco without whom this data would not be available. Finally, I would like to thank José Ignacio Hualde, Erik Willis, and three anonymous reviewers for their helpful insight into this topic. I have attempted to address their comments regarding this chapter. Any errors or omissions, however, remain my own. Support for this research project has been received from the University of Illinois at Urbana-Champaign through the Center for Latin American and Caribbean Studies (CLACS) Summer Field Research Grant and the Graduate College Dissertation Travel Grant.

<sup>1</sup> See Hualde (2002:103) and Willis (2003:15) for an extensive listing of research on Spanish peak alignment.

2002). The importance of peak placement is also highlighted in the discussion by Ladd (1996) of intonation across languages. Ladd (1996:129) notes that the phonetic misalignment of tunes may actually lead to phonological misinterpretation of stress. In cross-language communication, difficulty may be encountered if speakers of one language, such as Italian, are accustomed to early peak alignment, while others, such as English and German speakers, tend to place the peak at the end of the syllable, even when speaking Italian.<sup>2</sup>

Cross-linguistic influence in the placement of peaks can likewise be examined for certain varieties of Spanish, such as those found in Peru. The predominant indigenous language spoken in Peru is Quechua, which has over 3 million speakers spread across several subvarieties (Chirinos Rivera 2001). According to the 1993 census statistics reported in *Atlas lingüístico del Perú* (Chirinos Rivera), 63% of the population over the age of 5 is shown to speak Quechua in the Department of Cusco; in the Department of Lima, this percentage drops to less than 10%. Although Spanish and Quechua are genetically unrelated languages, their coexistence for over 500 years, coupled with extralinguistic factors such as relative prestige and political status, has given rise to observable language contact phenomena in the lexicon, phonology, morphology, syntax and pragmatics of both Spanish and Quechua (Escobar 1978; Escobar 1990, 1994, 2000; Godennzi 1996; Granda Gutiérrez 1995, 1999, 2001; Klee 1990, 1996). The extent to which the intonation system for Spanish as spoken in Peru has undergone a similar mutual influence is not yet known.

Descriptions of other Latin American Spanish varieties have often included the impressionistic claim that the intonation of indigenous languages has been borrowed by the local Spanish variety. For example, the intonation in central Mexico has been attributed to Nahuatl origins: "The intonation, in the general population, is identical to that employed in Nahuatl; in the educated classes, this local feature is attenuated. The final cadence of a declarative phrase is characteristically very different from the usual cadence of Castile" (Henríquez

---

<sup>2</sup> In a perception study of lexical stress in Castilian Spanish, the manipulation of peak placement alone was not enough to lead listeners to identify the syllable containing the peak as stressed (Llisterri et al. 2002). This result suggested that the perception of other correlates of stress, including duration and intensity, may need to be manipulated in combination with F0 maxima in order to ascertain how stress is perceived in Spanish. However, as noted by an anonymous reviewer, the study by Enríquez, Casado, and Santos (1989) found the fundamental frequency to be the most important factor in the perception of stress in Spanish, underlining the need for investigations on how intonation may be affected in situations of language contact.

Ureña 1938:335, my translation).<sup>3</sup> For Chilean Spanish, Amado Alonso refuted Rudolfo Lenz' thesis regarding Araucan influence almost categorically, yet still offered intonation as a viable point of language contact: "It is not necessary to discard the probability that Araucan, as a substratum language, or as an adstratum language, may have left some trace within Chilean Spanish, above all in the melodies and plays of rhythm" (Alonso 1940:289, my translation).<sup>4</sup> Sosa (1999:241-245) also provides an extended summary of claims of contact between indigenous languages and Latin American Spanish intonation.

In this chapter, to address this possibility of indigenous language influence, I report the findings from the instrumental analysis of two regional varieties of Spanish as spoken in Peru. The specific intonation feature I have analyzed is peak alignment. The first variety of Spanish is spoken in the coastal city of Lima and has historically had relatively little contact with Quechua (although recent waves of migration to the capital may have altered this reality somewhat); the second variety, Andean Spanish, is found in Cusco and other interior highland areas, and has experienced a more prolonged and intense contact with Quechua. The distinction between these two varieties was identified by Escobar (1978). In his analysis of sociolinguistic variation of Spanish within Peru, Escobar notes that Andean Spanish, among other characteristics, maintains the distinction between voiced palatal phonemes: The lateral approximant /*l*/ is distinguished from the fricative /*j*/, whereas Lima Spanish shows a fusion toward the fricative /*j*/. In the survey of indigenous languages and their role in the development of Latin American Spanish, Granda Gutiérrez (1995) cites the Andean region, Paraguay, and the Yucatán as the three most likely candidates for influence from indigenous languages, given the historical, political, and social conditions present in these areas.

Prior to outlining the current research study on Spanish, a very brief characterization of Quechua intonation peaks will be provided to highlight differences from the description of Spanish intonation previously stated. As Cerrón-Palomino (1987:128) notes, very little is currently known about Quechua intonation:

---

<sup>3</sup> Original: "La entonación, en las clases populares, es idéntica a la que se emplea al hablar náhuatl; en las clases cultas, el matiz local se atenúa. Es característica la cadencia final de la frase enunciativa, muy distinta de la cadencia usual en Castilla" (Henríquez Ureña 1938:335).

<sup>4</sup> Original: "No hay que descartar la probabilidad de que el araucano, ya como sustrato, ya como adstrato, haya dejado alguna huella en el chileno, sobre todo en las melodías y en los juegos rítmicos" (Alonso 1940:289).

The phenomena of accent, rhythm, and intonation are the least understood points within Quechua phonology. In this sense, the existing studies do not provide sufficiently exhaustive data from which one can postulate the originating suprasegmental features. Of the prosodic elements mentioned, only stress has received greater attention, due to its relatively discrete character. (my translation)<sup>5</sup>

Quechua is an agglutinating language in which stress is placed on the penultimate syllable of a word in the majority of the cases, although some words may bear final stress (Calvo Pérez 1993:49). Quechua behaves like an intonation language, using intonation to indicate changes in pragmatic meaning: a higher peak on the final word may be used to distinguish between simple declarative utterances and imperatives (Semanez Flórez 1996:111). In his characterization of Quechua intonation, Semanez Flórez adopts the terminology put forth by Navarro Tomás (1944) for Spanish. Specifically, utterance contours are distinguished according to the type of 'toneme' employed, which refers to the last lexically stressed syllable in an utterance along with the following boundary tones. In that sense, the final peak is presumed to be nuclear, as in Spanish. No information is given regarding peak alignment or other peaks within the utterance. Since more research is needed to establish the status of the final peak as nuclear in Quechua, the terminology 'prenuclear'-'nuclear' will be employed here only provisionally, and will be coupled with the more neutral terms of 'nonfinal' and 'final.'

My description of Quechua intonation is based on the instrumental analysis of recordings of two male Quechua speakers, which include over 100 utterances of both read and spontaneous speech. The finding on peak alignment as presented here is preliminary; currently, a more in-depth description of Quechua intonation features is underway. In the utterances analyzed thus far, both final (nuclear) and nonfinal (prenuclear) peaks in neutral declaratives are realized within the stressed syllable. In this case, peak alignment within the tonic syllable does not serve to indicate differences in final versus nonfinal position. It should be noted that these nonfinal peaks appearing within the tonic syllable do not necessarily indicate any particular highlighting through focus (see Figure 1).

---

<sup>5</sup> Original: "Los fenómenos de acento, ritmo y entonación son los puntos menos comprendidos dentro de la fonología quechua. En este sentido, los trabajos descriptivos existentes no proporcionan datos suficientemente exhaustivos a partir de los cuales se puedan postular los rasgos suprasegmentales originarios. De los elementos prosódicos mencionados, sólo el acento de intensidad ha recibido una mayor atención debido a su carácter relativamente discreto" (Cerrón-Palomino 1987:128).

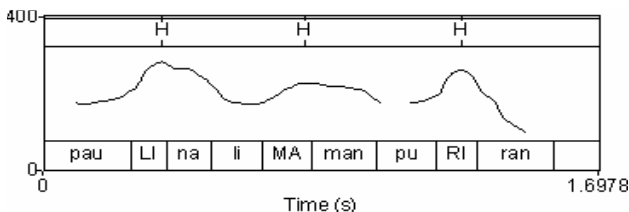


Fig. 1: *Quechua* utterance read by native *Quechua* speaker (Paulina Limaman puriran “Paulina went to Lima”; lexically stressed syllables are underlined)

Several proposals have been made regarding the phonological representation of the prenuclear and nuclear pitch accent in Spanish. Within the autosegmental-metrical approach (Pierrehumbert 1980), a series of high and low pitch targets are associated locally with specific stressed syllables. The combination of these highs and lows serves to form phonologically contrastive pitch accents that may be employed to express distinctions in meaning. Since the data presented here include only measurement of peak placement, and do not address the position of valleys, I will refrain from proposing a pitch-accent analysis for Peruvian Spanish at this time. However, I do refer the reader to the ongoing discussion of this area within the literature, including analysis of the factors affecting the alignment of peaks (e.g., Llisterri et al. 1995; Prieto, van Santen, & Hirshberg 1995), the adoption of the Sp\_ToBI annotation system for Spanish (Beckman et al. 2002), and recent reviews by Hualde (2002), McGory & Díaz-Campos (2002), Face (2002b), and Willis (2003).

In this study, the placement of peaks in Peruvian Spanish is examined in order to determine the extent to which Spanish as spoken in Lima and Cusco are similar to descriptions of other varieties of Spanish found in the literature. A comparison is made between Lima and Cusco varieties to explore regional differences in peak alignment. Additionally, within the Cusco variety, each participant’s knowledge of Quechua is noted in order to observe how Quechua may have exercised an influence on the development of distinct intonation patterns in the Andean region.

### 1. *Experiment*

Each participant read aloud a series of question-and-answer pairs printed on 72 index cards. The questions were designed to present different types of focal conditions or contexts in which the answers would be produced. The data set presented here includes only the responses to the broad-focus questions, which do not emphasize any one particular lexical item over another, for example, “What is happening?” and “John is eating dinner.” The set of index cards were

ordered using a spreadsheet randomizer so that the same focus conditions were not immediately adjacent. Participants read the cards in the same sequence, resulting in a pseudo-randomization of the test materials. The cards were divided into two even blocks with 12 distractors placed within the test materials, including two distractors at the beginning and end of each block. Each block of cards was read with a break in between. Then the cards were read in reverse order, also with breaking between blocks. A total of 24 broad-focus productions were recorded per participant (12 broad-focus utterances x 2 repetitions).

The actual structure of the target response always appeared in the order subject, verb, object (SVO), which is considered a neutral word order for Spanish sentences with transitive verbs. (Broad-focus questions and responses can be found in Table A1 of the appendix.) The target words themselves included both open and closed syllables as well as words with final, penultimate, and antepenultimate stress. Llisterra et al (1995) found that the alignment of peaks varied according to lexical stress, so that peaks appeared on the post-tonic syllable more frequently in words with penultimate stress than with other stress patterns. For this study, only open syllables with penultimate stress are analyzed, in order to not mix syllable types and stress patterns. Also, both Quechua and Spanish contain open syllables and have penultimate stress as the most predominant stress pattern. Therefore, the findings from this study of Spanish can later be compared to other future studies of Quechua. For each of the measurements taken, there were at least two syllables between the stressed vowel being measured and the subsequent lexically stressed vowel. Voiced consonants were used in the target words to ensure a continuous pitch contour, with the exception of the plural *-s*. Other voiceless consonants that do appear are not adjacent to the stressed syllable being measured.

### 1.1 *Recordings*

Participants spoke into a Shure 512 headset microphone that was positioned within an inch of the mouth. The recording was made with a Sharp MD-SR60 minidisc recorder using Sony minidisks. The recording was then transferred to the computer and analyzed with Praat 4.0.47, using autocorrelation of the fundamental frequency (F0). In Lima, the recordings were made at the Pontificia Universidad Católica del Perú (PUCP). Recordings in Cusco were made at the Centro Bartolomé de las Casas (CBC) Escuela Andina de Postgrado. Speakers received a nominal remuneration for their participation in the study, which lasted approximately 45 minutes. In addition to the oral reading task, each participant completed a language history questionnaire that

detailed their experience with Spanish and Quechua, including their interaction with family and community members.

### 1.2 Data measurement

For the 12 broad-focus utterances, the total number of possible measurements of peaks associated with open syllables with penultimate stress is 38 per person: (5 for subject position + 8 for verb position + 6 for object position = 19 x 2 repetitions). The fundamental frequency (F0) maximum was measured in milliseconds (ms) with respect to the end of the stressed syllable. If the peak is realized after the stressed syllable, the temporal placement of the peak is positive. If the peak is realized before the end of the stressed syllable, the peak-placement values are negative, as shown in Figure 2.

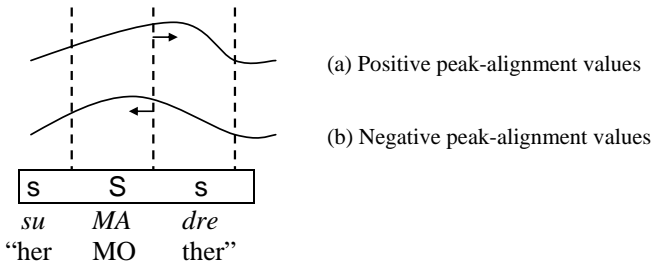


Fig. 2: Measurement of peak alignment relative to the offset of the stressed syllable (S)

The participants were instructed to read the cards as naturally as possible. They were told that they needed to read the question and answer separately, rather than as one long utterance. However, in some cases, speakers also produced pauses within the utterance. Since analyzing the effect of pauses on peak alignment is not the aim of this study, measurements were not taken of the subject or the verb if a pause occurred after either of these. A measurement was still taken of the object in nuclear position, since the final stressed syllable is always followed by a pause. In other cases, if the boundary between syllables was unclear, a measurement was not recorded. Last, a reading was not taken if no peak was apparent, that is, if there was no appreciable rise greater than 7 Hz to the peak, or no appreciable fall greater than 7 Hz to a valley (F0 minimum). The value of 7 Hz was chosen as a threshold, but may be considered somewhat arbitrary in that studies are needed to determine if peaks may be perceived with a lesser difference in surrounding tonal target heights or if an even greater difference is necessary. This operational definition of 7 Hz is



used also to be comparable to other studies (e.g., Willis 2002).<sup>6</sup> The resulting data set represents only those peaks that were measurable according to the preceding criteria. A minimum of five measurable peaks each for subject, verb, and object were needed per speaker in order to be included in the data set presented here.

### 1.3 *Participants*

Twenty speakers of Peruvian Spanish were recorded for the experiment. Participants in this study have been divided according to their origin and their knowledge of Spanish and Quechua. The speakers in Lima were all monolingual Spanish speakers whose parents and grandparents were likewise monolingual Spanish speakers from the city of Lima (speaker codes L01-L05). Cusco participants were all living in the city of Cusco at the time, and were either from the city or at least from the larger Department of Cusco. Participants whose parents were not from the Department of Cusco are not included in this data set. All Cusco speakers reported having parents who spoke Quechua to some degree, as well as Spanish. Only one participant's parents spoke only Quechua. The participants themselves can be divided into three groups: those who reported speaking only Spanish before beginning school at approximately age 5 (speakers C01-C07), those who spoke both Quechua and Spanish (C21-C25), and those who spoke only Quechua in the home before entering the school system (C31-C33).

Three remaining descriptive characteristics that are held in common between the two groups are: gender, age range, and level of education. First, all speakers from both Lima and Cusco were male, ranging in age from 18 to 39. All Lima participants were university students of the PUCP. Likewise, participants in Cusco were enrolled in or had received postsecondary education. (See Table A3 in the appendix for a more detailed description of each participant.)

---

<sup>6</sup> As noted by an anonymous reviewer, the peaks that do not show a rise of less than 7 Hz may actually still represent valid tonal patterns within these Spanish varieties. Therefore, the findings here are presented with the understanding that other, less frequent patterns may also be present. The percentage of measurements that did not meet the previous criteria is relatively low: In the subject position, all cases showed a rise. For the Lima group, 22% of the stressed syllables in verb position and 19% in object position lacked an appreciable rise; the Cusco group showed 11% in verb position and 13% in object position.

## 2. Results and discussion

### 2.1 Lima peak placement

The data collected for the Lima group show that prenuclear peaks associated with open syllables with penultimate stress are almost exclusively realized in the post-tonic syllable for both the initial peak (subject) and medial peak (verb). These may be termed ‘late’ peaks with respect to the end of the stressed syllable. On the other hand, the peak associated with the object in final, nuclear position is realized within the tonic syllable. These peaks are considered to be ‘early’ since they appear before the offset of the stressed syllable (see example in Figure 3). The peaks are labeled with an ‘H,’ corresponding to the local F0 maxima (as opposed to a phonological representation of a pitch accent H\*, L\*+H, etc.).

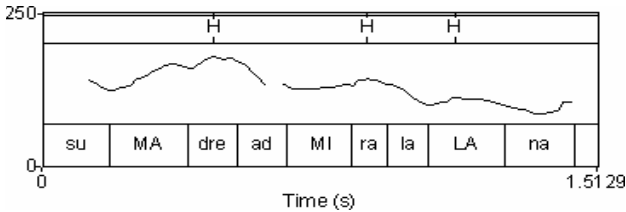


Fig. 3: Late initial and medial peaks and early final peak for the phrase Su madre admira la na “Her mother admires the wool” (Participant L01; lexically stressed syllables underlined)

The box plots in Figure 4 give a summary of the results for each speaker. These box plots represent the 25th through 75th percentile range of data collected for each individual. The bar within the box represents the median. The lines extending from the box indicate the limits of the 10<sup>th</sup> and 90<sup>th</sup> percentile range. Any observation appearing outside these extremes is represented with a circle. Each box plot contains a minimum of five observations for subject, verb, and object positions. (See Tables A2 and A4 in the appendix for the count of data points used to calculate the box plots for each speaker.)

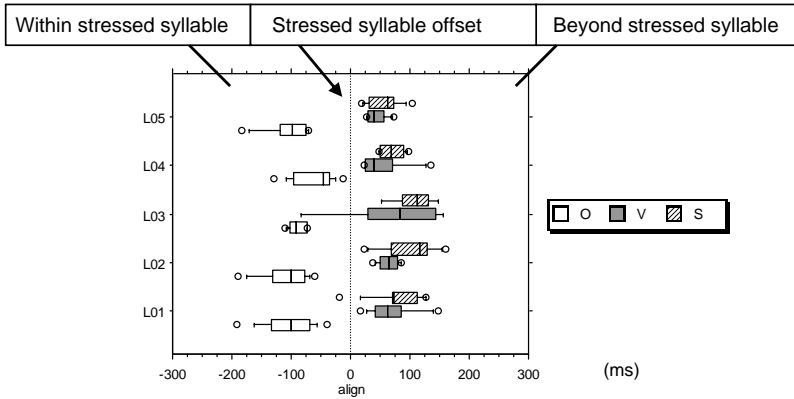


Fig. 4: Lima peak alignment in (ms) with respect to stressed syllable boundary ('O' = nuclear object, 'V' = prenuclear verb, 'S' = prenuclear subject)

The data collected from the group of speakers in Lima demonstrate a placement of prenuclear peaks that is consistent with previous reports for other varieties of Peninsular and Latin American Spanish. In Garrido et al. (1995), Peninsular Spanish showed 70% of F0 maxima appearing after the stressed syllable. Mota Gorriz (1997) also found F0 peaks placed on the post-tonic syllable in over 80% of the cases for Peninsular Spanish. In his analysis of Madrid Spanish, Face (2002b:21) found prenuclear words in medial position within an utterance to appear on average 52 ms ( $N = 290$ ,  $SE = 3.5$ ) after the end of the stressed syllable. For Mexican Spanish as spoken in Puebla, Willis (2002) notes that prenuclear peaks are most frequently aligned late in the neutral declaratives. Willis (2003) also notes for Dominican Spanish that there is a wide range of post-tonic high tone alignments for both broad and contrastive focus, with a global mean of 102 ms ( $N = 232$ ,  $SD = 56$ ). To provide points of comparison with other studies, the global mean of prenuclear peak placement for the Lima speakers is 84 ms ( $SD = 35$ ) for the subject position and 62 ms ( $SD = 42$ ) for the verb (see Table A2 in the appendix for a listing of means for each individual speaker). These values are in between those found for Madrid Spanish and Dominican Spanish. Moreover, there is only one instance (with participant L03) of early peak alignment on the verb; the remaining four participants from Lima show all prenuclear peaks with late alignment. In relative terms, the peak is realized on average 57% into the duration of the post-tonic syllable in initial subject position, and 56% in medial verb position. A one-way ANOVA was conducted for the Lima group, with

alignment of the peak as the dependent variable and position within the utterance as the independent variable. A significant effect was found for position ( $F(2,133) = 260.02$ ,  $p < 0.001$ ). A Tukey post hoc analysis shows significant differences at the 0.05 level between the nuclear object and each of the prenuclear peaks, but not between the prenuclear subject and verb. These results show that Lima speakers demonstrate the same trend as other Spanish varieties of realizing prenuclear peaks after the stressed syllable boundary.

## 2.2 *Cusco peak placement*

For the group of Cusco speakers, there appear to be different patterns of peak alignment. The data for each speaker are arranged according to two criteria: (a) whether the mean alignment value is considered 'early' (within the tonic syllable) or 'late' (after the tonic syllable), and (b) whether the box plot itself extends to within the tonic syllable or into the post-tonic (i.e., if the values between the 25<sup>th</sup> and 75<sup>th</sup> percentiles can be considered 'early' or 'late'). From these criteria, four patterns of peak alignment have been identified. However, these patterns may actually be considered as slices along a continuum, such that other valid divisions may likewise be possible. These groupings have been made in order to begin to describe similarities between the behaviors observed. The first pattern will be labeled Pattern A. Similar to the peak-alignment distribution found for Lima and other descriptions of Spanish previously mentioned, prenuclear peaks in Pattern A are realized 'late,' on the post-tonic syllable (for both subject and verb), and the nuclear peak is realized 'early,' within the tonic syllable (see box plots in Figure 5). Pattern B likewise shows the same general trend, although there are incidences of prenuclear peaks within the tonic syllable. However, the mean values are still beyond the stressed syllable for all speakers included in this group (see box plots in Figure 6). The last two patterns, C and D, are notably different from the previous two. In Pattern C, the mean of the prenuclear verb peaks appears within the stressed syllable, and in Pattern D, the mean values for both prenuclear subject and verb are realized within the tonic syllable (see examples in Figures 7-8; box plots for individual speakers appear in Figures 9-10). In these last two patterns, prenuclear peaks are not distinguished from nuclear peaks with respect to tonic syllable boundary since both may be realized within the stressed syllable. A summary of the four patterns described is listed in Table 1.

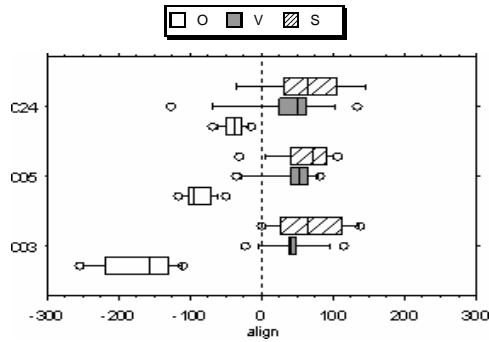


Fig. 5: Cusco peak alignment (ms), Pattern A

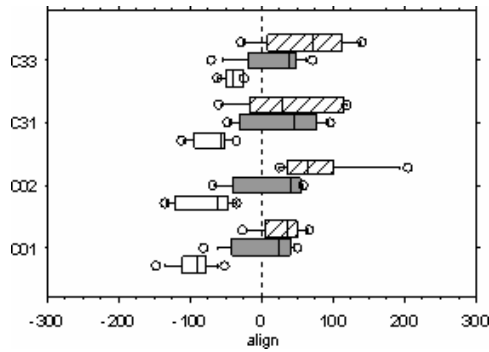


Fig. 6: Cusco peak alignment (ms), Pattern B

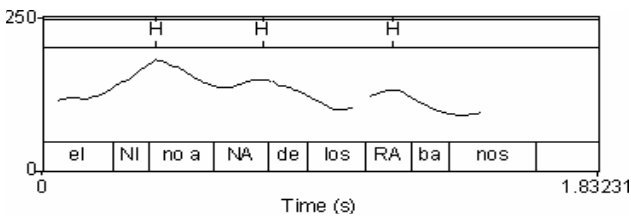


Fig. 7: Late initial peak, early medial and final peaks for the phrase *El niño añade los rábanos* “The boy adds the radishes” (Participant C22; lexically stressed syllables underlined)

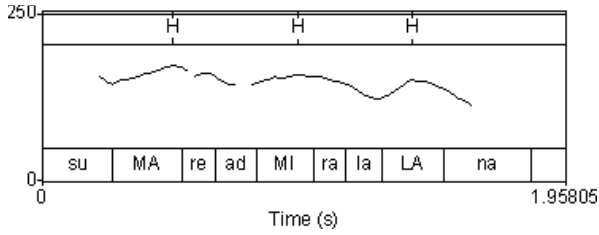


Fig. 8: Early initial, medial, and final peaks for the phrase Su madre admira la lana “His mother admires the wool” (Participant C07; lexically stressed syllables underlined)

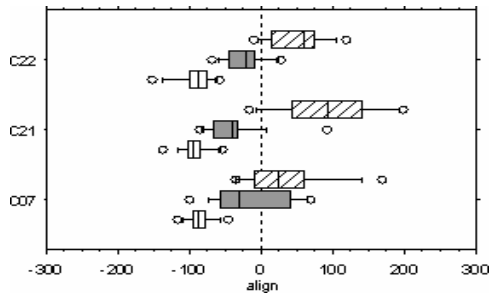


Fig. 9: Cusco peak alignment (ms), Pattern C

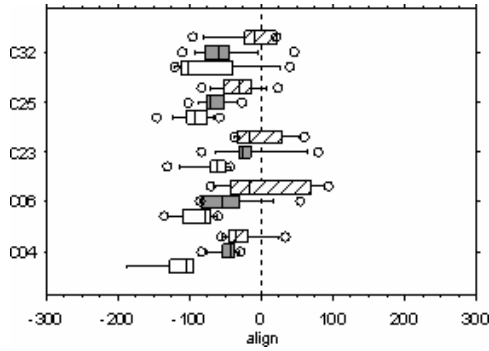


Fig. 10: Cusco peak alignment (ms), Pattern D

Peak-placement groupings	Initial peak (subject position)			Medial peak (verb position)		
	Mean	Percentiles		Mean	Percentiles	
		25 <sup>th</sup>	75 <sup>th</sup>		25 <sup>th</sup>	75 <sup>th</sup>
Pattern A	L	L	L	L	L	L
Pattern B	L	L,E	L	L	E	L
Pattern C	L	L,E	L	E	E	L,E
Pattern D	E	E	L	E	E	E

Table 1: *Patterns of Cusco peak alignment, according to alignment with stressed syllable ('E' = early, within tonic syllable; 'L' = late, on post-tonic syllable)*

In order to compare these results for Cusco with those described earlier for Lima and other varieties of Spanish, a summary of global means is included in Table 2 (see also Table A4 in the appendix for global means of individual speakers). The Cusco speakers show similar (although slightly lower) subject means compared to Lima in Patterns A through C and verb means in Pattern A; however, the mean peak placement for the verb in Pattern B is strikingly closer to the stressed syllable offset (17 ms). The remainder of the means are negative for the subjects in Pattern D, and for the verb in Patterns C and D, such that peaks are aligned within the stressed syllable. These negative average peak-placement values are not found for Lima Spanish speakers in this study and are not reported for prenuclear peaks in Madrid Spanish and Dominican Spanish as discussed earlier.

Origin	Pattern	Number of speakers	Initial peak (S)			Medial peak (V)		
			<i>n</i>	Mean (ms)	<i>SD</i>	<i>n</i>	Mean (ms)	<i>SD</i>
Lima		5	36	84	(35)	48	62	(42)
Cusco	Pattern A	3	21	65	(52)	28	40	(48)
Cusco	Pattern B	4	27	52	(58)	34	17	(51)
Cusco	Pattern C	3	25	58	(60)	39	-25	(43)
Cusco	Pattern D	5	41	-14	(40)	51	-46	(33)

Table 2: *Global means in (ms) for peak placement with respect to stressed syllable boundary*

For each subsequent pattern identified, there are fewer late peaks appearing in prenuclear position. As seen in Table 3, Patterns A and B are similar to those described earlier for Peninsular Spanish (70%-80%). However, there is a sharp decrease in subject late peaks to less than 50% in Pattern D; for verbs, instances of late peaks drop to less than 35% in Pattern C and 20% or less in Pattern D. To test the grouping of specific speakers into the patterns described, a mixed-model ANOVA was conducted for the Cusco group with peak alignment taken as the dependent variable, pattern as the fixed factor, and person nested within pattern as the random factor.

	# Initial peak (S)	% Late peaks	# Medial peak (V)	% Late peaks
Pattern A				
C03	6/6	100%	7/8	88%
C05	9/10	90%	9/11	82%
C24	4/5	80%	8/10	80%
Pattern B				
C01	7/9	78%	7/11	64%
C02	6/6	100%	5/7	71%
C31	3/6	50%	4/6	67%
C33	5/6	83%	7/10	70%
Pattern C				
C07	6/8	75%	4/12	33%
C21	7/8	88%	1/13	8%
C22	8/9	89%	3/14	21%
Pattern D				
C04	1/7	14%	0/8	0%
C06	4/9	44%	1/10	10%
C23	3/8	38%	2/10	20%
C25	1/10	10%	0/11	0%
C32	3/7	43%	1/12	8%

Table 3: *Instances of late peaks according to pattern for Cusco speakers*

From the pairwise comparisons shown in Table 4, it appears that two patterns are needed to distinguish between the initial peak: (ABC) and D. Three patterns are needed to distinguish between the medial peak: A, B, and (CD), although C and D approach significance also. Since these groupings do not overlap, four patterns have been identified overall to describe the different alignment patterns found. However, more speakers and more tokens per speaker are needed to further examine the alignment patterns and the assignment of speakers into these patterns as described here.

Pattern	Initial peak (S): $p$	Mean difference
A D	< 0.001	78.404
B D	< 0.001	63.493
C D	< 0.001	72.025
Pattern	Medial peak (V): $p$	Mean difference
A B	0.048	25.319
A C	< 0.001	65.701
A D	< 0.001	86.658
B C	0.004	40.382
B D	< 0.001	61.339
C D	0.063	20.957

Table 4: *Pairwise comparison of patterns from mixed-model ANOVA*



### 2.3 *Peak-placement patterns and the question of influence from Quechua*

The placement of peaks for the Lima group includes data only from monolingual Spanish speakers with no personal knowledge of Quechua. As stated, the post-tonic alignment of prenuclear peaks in this group coincides with other descriptions of Peninsular and Latin American Spanish varieties. In the Cusco group, the four patterns do not coincide with the speakers' knowledge of Quechua. Participants who spoke only Spanish before starting school (C01-C07) demonstrate all four patterns of peak placement. Those who spoke both Quechua and Spanish (C21-C25) demonstrate Pattern A with late prenuclear peaks and Patterns C and D with early prenuclear peaks. Finally, those who spoke only Quechua (C31-C33) before starting school demonstrate Pattern B with later prenuclear peaks and Pattern D with early prenuclear peaks. Yet, as a whole, what is most remarkable with this data set is that the majority of the Cusco speakers demonstrate patterns of peak alignment (Patterns B-D) other than those reported for Lima and other Spanish varieties. For these Cusco speakers, prenuclear peaks are closer to the stressed syllable offset, if not entirely within the stressed syllable.

One possible explanation for the peak alignments found in this data set may be that the language contact situation between Quechua and Spanish in Cusco has given rise to alternate peak-alignment patterns not found outside the contact region. The fact that Lima Spanish does not demonstrate early prenuclear peak alignment supports this claim. Elordieta (2003) also found tonic alignment of prenuclear peaks in another variety, Lekeitio Spanish, which is in contact with Northern Bizkaian Basque. In an examination of the intonation of Turkish-German bilinguals, Queen (1996) suggests that fluent bilingual speakers may actually create a mixed system that draws from both languages, such that the resulting intonation patterns employed may differ from those used by monolinguals of either language. Sosa (1999:187-188) indicates that prenuclear peaks in Buenos Aires Spanish, a variety historically in contact with Italian, are aligned with the stressed syllable. As more research is conducted on other contact and noncontact varieties of Spanish, the extent to which tonic alignment of prenuclear peaks can be attributed to outside linguistic influence will become more apparent. A feature shared by all Cusco speakers in this study is exposure in varying degrees to Quechua and to Quechua-influenced Spanish, through personal knowledge of the indigenous language, communication with Quechua-speaking family members, or interaction within the community. Therefore, it is plausible that contact with Quechua, either direct or indirect, may have contributed to the increased

realization of prenuclear peaks within the stressed syllable in the Spanish spoken in Cusco.

### 3. *Summary*

This chapter offers an analysis of two regional varieties of Peruvian Spanish with respect to one feature of intonation: the placement of peaks in prenuclear position. It was found that Lima speakers and some Cusco speakers follow the trait observed in other Spanish varieties of realizing prenuclear peaks on the post-tonic syllable in open syllables within broad-focus declaratives. However, several speakers from Cusco demonstrated alignment of prenuclear peaks (subject and/or verb) within the stressed syllable. It has been suggested that the long-standing language contact situation with Quechua may have resulted in the development of these alternate intonation patterns. Other varieties of Spanish, such as Madrid Spanish, have been reported to align peaks within the stressed syllable under contrastive focus. In order to gain a fuller understanding of the intonation system employed by these Cusco speakers, it will be necessary to next ascertain how contrastive focus is realized, since these Cusco speakers already produce early prenuclear peaks under broad-focus conditions. Further analysis of Quechua intonation will likewise provide additional insight into this indigenous language and the ways it may contribute to the development of Spanish intonation as found in the Andean region.

## APPENDIX

<i>¿Qué pasa?</i> “What is happening?”	<i>¿Qué pasaba?</i> “What was happening?”	<i>¿Qué pasará?</i> “What will happen?”
<i>El <u>niño</u> <u>añade</u> los rábanos.</i> “The child adds the radishes.”	<i>Amalia <u>podaba</u> los árboles.</i> “Amalia pruned the trees.”	<i>Su <u>familia</u> <u>mandará</u> los <u>violines</u>.</i> “His family will send the violins.”
<i>Yolanda <u>domina</u> el <u>castellano</u>.</i> “Yolanda knows Spanish.”	<i>El criminal <u>llevaba</u> el <u>ídolo</u>.</i> “The criminal carried the idol.”	<i>Bernardo <u>venderá</u> los <u>mangos</u>.</i> “Bernardo will sell the mangos.”
<i>El <u>vándalo</u> <u>agarra</u> los <u>baldes</u>.</i> “The vandal grabs the jugs.”	<i>El <u>águila</u> <u>guardaba</u> el <u>nido</u>.</i> “The eagle guarded the nest.”	<i>El <u>albañil</u> <u>moverá</u> los <u>barriles</u>.</i> “The mason will move the barrels.”
<i>Su <u>madre</u> <u>admira</u> la <u>lana</u>.</i> “Her mother admires the wool.”	<i>La <u>víbora</u> <u>devoraba</u> los <u>animales</u>.</i> “The snake devoured the animals.”	<i>Su <u>hermana</u> <u>retirará</u> la <u>demanda</u>.</i> “Her sister will withdraw the complaint.”

Table A1: *Broad-focus questions and answers; the stressed syllable is underlined in the words with open syllables with penultimate stress*

Speakers	Prenuclear subject			Prenuclear verb			Nuclear object		
	<i>N</i>	Mean (ms)	<i>SD</i>	<i>N</i>	Mean (ms)	<i>SD</i>	<i>N</i>	Mean (ms)	<i>SD</i>
L01	9	81	(44)	15	71	(39)	12	-105	(43)
L02	7	103	(47)	8	64	(17)	11	-109	(41)
L03	5	108	(36)	5	73	(95)	10	-91	(14)
L04	6	70	(21)	10	55	(41)	12	-60	(36)
L05	9	56	(28)	10	45	(17)	7	-106	(40)
Avg		84	(35)		62	(42)		-94	(35)

Table A2: *Peak-placement means (in ms) with respect to stressed syllable offset for Lima*

Code	Origin	Age range	Education* : stage : years	Language before	Parents'
				5 years	language(s)
L01	Lima	25-29	PUCP : completed : n.a.	Spanish	Spanish
L02	Lima	20-24	PUCP : in progress : 2	Spanish	Spanish
L03	Lima	20-24	PUCP : in progress : 2	Spanish	Spanish
L04	Lima	25-29	PUCP : in progress : 5	Spanish	Spanish
L05	Lima	20-24	PUCP : in progress : 3	Spanish	Spanish
C01	Cusco	20-24	ISMLAM : in progress : 2	Spanish	Qu, Sp
C02	Cusco	18-19	ISMLAM : in progress : 1	Spanish	Qu, Sp
C03	Cusco	30-34	UNSAAC : in progress : n.a.	Spanish	Qu, Sp
C04	Cusco	18-19	ISMLAM : in progress : 2	Spanish	Qu, Sp
C05	Cusco	30-34	Postsecondary : completed : n.a.	Spanish	Qu, Sp
C06	Cusco	25-29	Universidad Andina : n.a. : n.a.	Spanish	Qu, Sp
C07	Cusco	20-24	ISMLAM : in progress : 2	Spanish	Qu, Sp
C21	Cusco	20-24	UNSAAC : completed : n.a.	Qu, Sp	Qu, Sp
C22	Cusco	20-24	UNSAAC : in progress : 2	Qu, Sp	Qu, Sp
C23	Cusco	30-34	UNSAAC : progress : 4	Qu, Sp	Qu, Sp
C24	Cusco	20-24	UNSAAC : in progress : 4	Qu, Sp	Qu, Sp
C25	Cusco	35-39	UNSAAC : completed : 5	Qu, Sp	Qu, Sp
C31	Cusco	30-34	UNSAAC : completed : n.a.	Quechua	Qu, Sp
C32	Cusco	25-29	Postsecondary : completed : n.a.	Quechua	Qu, Sp
C33	Cusco	25-29	Postsecondary : completed : n.a.	Quechua	Quechua

\*PUCP = Pontificia Universidad Católica del Perú ; UNSAAC = Universidad Nacional San Antonio Abad del Cusco ; ISMLAM = Instituto Superior de Música "Leandro Alviña Miranda"; Postsecondary = Studies beyond high school, institution not specified; n.a. = not provided. If completed, 3-5 years minimum assumed.

Table A3: *Description of participants*

Pattern	Speaker	Prenuclear subject		Prenuclear verb			Nuclear object			
		<i>N</i>	Mean (ms)	<i>SD</i>	<i>N</i>	Mean (ms)	<i>SD</i>	<i>N</i>	Mean (ms)	<i>SD</i>
A	C03	6	68	(51)	8	43	(37)	7	-172	(54)
A	C05	10	63	(40)	10	42	(40)	12	-88	(20)
A	C24	5	64	(66)	10	35	(67)	10	-39	(16)
	Avg		65	(52)		40	(48)		-100	(30)
B	C01	9	26	(32)	11	2	(45)	9	-95	(27)
B	C02	6	82	(65)	7	14	(54)	9	-82	(41)
B	C31	6	36	(73)	6	31	(60)	8	-69	(27)
B	C33	6	63	(63)	10	19	(46)	11	-41	(14)
	Avg		52	(58)		17	(51)		-72	(27)
C	C07	8	34	(66)	12	-15	(55)	12	-85	(19)
C	C21	8	93	(73)	13	-39	(45)	11	-92	(23)
C	C22	9	48	(42)	14	-22	(30)	10	-92	(28)
	Avg		58	(60)		-25	(43)		-90	(23)
D	C04	7	-27	(30)	8	-48	(17)	5	-118	(39)
D	C06	9	4	(62)	10	-47	(42)	7	-89	(27)
D	C23	8	-3	(37)	10	-15	(47)	9	-68	(28)
D	C25	10	-32	(29)	11	-65	(21)	12	-93	(23)
D	C32	7	-12	(41)	12	-56	(39)	7	-73	(59)
	Avg		-14	(40)		-46	(33)		-88	(35)

Table A4: *Peak-placement means (in ms) with respect to stressed syllable offset for Cusco*

## REFERENCES

- Alonso, Amado. 1940. "La interpretación araucana de Lenz para la pronunciación chilena". *El español en Chile*. ed. by Amado Alonso & Raimundo Lida, 279-289. Buenos Aires: Universidad de Buenos Aires
- Beckman, Mary E., Manuel Díaz-Campos, Julia T. McGory, & Terrell A. Morgan. 2002. "Intonation Across Spanish in the Tones and Break Indices Framework". *Probus* 14.9-36.
- Calvo Pérez, Julio. 1993. *Pragmática y gramática del quechua cuzqueño*. Cusco, Peru: Centro de Estudios Regionales Andinos Bartolomé de las Casas.
- Cerrón-Palomino, Rodolfo. 1987. *Lingüística Quechua*. Cusco, Peru: Centro de Estudios Rurales Andinos Bartolomé de las Casas.
- Chirinos Rivera, Andrés. 2001. *Atlas lingüístico del Perú (Estudios Urbanoregionales 6)*. Cusco, Peru: Ministerio de Educación, Centro de Estudios Regionales Bartolomé de las Casas.
- Elordieta, Gorka. 2003. "The Spanish Intonation of Speakers of a Basque Pitch-Accent Dialect". *Catalan Journal of Linguistics* 2.67-95.
- Enríquez, Emilia V., Celia Casado, & Andrés Santos. 1989. "La percepción del acento en español". *Lingüística Española Actual* 11.241-269.
- Escobar, Alberto. 1978. *Variaciones sociolingüísticas del castellano del Perú (Peru Problema 18)*. Lima: Instituto de Estudios Peruanos.
- Escobar, Anna María. 1990. *Los bilingües y el castellano en el Perú*. Lima: Instituto de Estudios Peruanos.
- , 1994. "Andean Spanish and Bilingual Spanish: Linguistic characteristics". *Language in the Andes* ed. by Bariella Hermon, Peter Cole, & Mario Daniel Martín, 51-73. Newark, DE: University of Delaware.
- , 2000. *Contacto social y lingüístico: El español en contacto con el quechua en el Perú*. Lima: Pontificia Universidad Católica del Perú.
- Face, Timothy L. 2002a. *Intonational Marking of Contrastive Focus in Madrid Spanish, LINGCOM Studies in Romance Linguistics*. Muenchen: Lincom.
- , 2002b. "Local Intonational Marking of Spanish Contrastive Focus". *Probus* 14.71-92.
- Garrido, Juan M., Joaquim Llisterri, Carme de la Mota, Rafael Marín, & Antonio Ríos. 1995. "Estudio comparado de las características prosódicas de la oración simple en español en dos modalidades de lectura". *Phonetica. Trabajos de fonética experimental* ed. by Angelika Braun & Herbert R. Masthoff, 173-194. Bilbao: Universidad de Deusto.
- Godenzzi, Juan Carlos. 1996. "Transferencias lingüísticas entre el quechua y el español". *Contactos y transferencias lingüísticas en Hispanoamérica (Signo y Señal 6)* ed. by Germán de Grandá Gutiérrez, 71-99. Buenos Aires: Universidad de Buenos Aires.
- Grandá Gutiérrez, Germán de. 1995. "El influjo de las lenguas indoamericanas sobre el español. Un modelo interpretativo sociohistórico de variantes areales de contacto lingüístico". *La lengua española y su expansión en la época de Tratado de Tordesillas* ed. by César Hernández Alonso, 99-117. Valladolid: Sociedad V Centenario del Tratado de Tordesillas.
- , 1999. *Español y lenguas indoamericanas en Hispanoamérica: Estructuras, situaciones y transferencias*. Valladolid, Spain: Universidad de Valladolid.
- , 2001. *Estudios de lingüística andina*. Lima: Pontificia Universidad Católica del Perú.
- Henríquez Ureña, Pedro. 1938. *El español en Méjico, los Estados Unidos, y la América Central*. Buenos Aires: Universidad de Buenos Aires.

- Hualde, José Ignacio. 2002. "Intonation in Spanish and the Other Ibero-Romance Languages: Overview and status quaestionis". *Romance Phonology and Variation. (Current Issues in Linguistics Theory)* ed. by Caroline Wiltshire & Joaquim Camps, 101-115. Amsterdam & Philadelphia: John Benjamins.
- Klee, Carol A. 1990. "Spanish-Quechua Language Contact: The clitic pronoun system in Andean Spanish". *Word* 41.35-46.
- , 1996. "The Spanish of the Peruvian Andes: The influence of Quechua on Spanish language structure". *Spanish in Contact* ed. by Ana Roca & John Jensen, 73-91. Somerville, MA: Cascadilla Press.
- Ladd, D. Robert. 1996. *Intonational Phonology*. Cambridge: Cambridge University Press.
- Llisterri, Joaquim, Rafael Marin, Carme de la Mota, & Antonio Ríos. 1995. "Factors Affecting F0 Peak Displacement in Spanish". Paper presented at ESCA 4th European Conference on Speech Communication and Technology, Madrid, September 1995.
- Llisterri, Joaquim, María Jesus Machuca, Carme de la Mota, Montserrat Riera, & Antonio Ríos. 2002. "The Role of F0 Peaks in the Identification of Lexical Stress in Spanish". *Phonetics and Its Applications. Festschrift for Jens-Peter Köster on the occasion of his 60th birthday (Zeitschrift für Dialektologie und Linguistik, Beiheft 121)* ed. by Angelika Braun & Herbert R. Masthoff, 350-361. Stuttgart: Steiner.
- McGory, Julia & Manuel Díaz-Campos. 2002. "Declarative Intonation Patterns in Multiple Varieties of Spanish". *Structure, Meaning, and Acquisition in Spanish: Papers from the 4th Hispanic Linguistic Symposium* ed. by James F. Lee, Kimberly L. Geeslin, & J. Clancy Clements, 73-92. Somerville, MA: Cascadilla Press.
- Mota Gorriz, Carme de la. 1997. "Prosody of Sentences With Contrastive New Information in Spanish". Paper presented at the ESCA Workshop on Intonation: Theory, Models, and Applications, Athens, September 1997.
- Navarro Tomás, Tomás. 1944. *Manual de pronunciación española*. New York: Hispanic Institute in the United States.
- Pierrehumbert, Janet. 1980. *The Phonology and Phonetics of English Intonation*. Ph.D. Dissertation, Massachusetts Institute of Technology.
- Praat: Doing phonetics by computer. Version 4.0.47. Paul Boersma & Paul Weenink, Amsterdam: University of Amsterdam. (Available online: <http://www.praat.org>)
- Prieto, Pilar, Jan van Santen, & Julia Hirshberg. 1995. "Tonal Alignment Patterns in Spanish". *Journal of Phonetics* 23.429-451.
- Queen, Robin Michelle. 1996. *Intonation in Context: A study of Turkish-German bilingual intonation patterns*. Ph.D. dissertation, University of Texas at Austin.
- Semanez Flórez, David I. 1996. *Gramática del quechua del qosqo*. Cusco, Peru: Academia Mayor de la Lengua Quechua.
- Sosa, Juan Manuel. 1999. *La entonación en español*. Madrid: Cátedra.
- Willis, Erik W. 2002. "Is There a Spanish Imperative Intonation Revisited: Local considerations". *Linguistics* 40.347-374.
- , 2003. *The Intonational System of Dominican Spanish: Findings and analysis*. Ph.D. Dissertation, University of Illinois at Urbana-Champaign.



# ON INTENTIONAL CAUSATION IN ITALIAN\*

ANTONELLA VECCHIATO  
*University of Southern California*

## 0. *Introduction*

One of the characteristic features of formal generative grammar is the existence within its explanatory apparatus of elements that are non-overtly realized. This article investigates apparently unrelated phenomena in Italian that call for the existence of a tacit intentional predicate at the semantic level. I assume a neo-Davidsonian framework arguing that predicates contain event arguments and sentences existentially quantify over events and event complexes formed of subevents. The first piece of data, at first observed in English, considers the modification with the adverb *quasi* “almost,” which creates ambiguous sentences when the event described is an intentional action (section 1). I claim that the ambiguity is the effect of the following combination of factors. First, the meaning of *quasi* is a binary predicate true of an event and an intensional entity (property or proposition). Second, the logical form of causative sentences with an intentional subject contains a tacit intentional predicate. The third and final factor is the semantic scope of the existential operator quantifying over events and the intension operator provided by the second argument of *quasi*.

The second phenomenon is a contrast between agent subjects and causer subjects with the *fare* periphrastic causative construction (make + VP) (section 2). Some conditions on the causal dynamics of the situation described by this causative construction hold only when the subject is intentional. These conditions are the result of the speaker’s fine-grained conception of the event triggered by the presence of a tacit intentional predicate. Similarly, the presence of an overt intentional predicate activates the conception of fine-grained causal dynamics affecting the entailments among sentences.

---

\* Extensive discussions with Jim Higginbotham and Barry Schein helped me considerably in developing the issues in this chapter. I would also like to thank for their comments Hagit Borer, José Camacho, Viviane Déprez, J.-R Hayashishita, Liliana Sánchez, Roumyana Pancheva, Juan Uriagereka, Jean-Roger Vergnaud, and three anonymous reviewers. Any mistakes are, however, solely mine. Some parts of this chapter were also discussed in my contribution to the Proceedings of WECOL (Vecchiato to appear).



Finally, the interpretation of the proform *lo* as referring to a predicate, when in a construction with *fare* (*farlo* “do so”), requires the contextual focus on the subject’s intention (section 3). This proform, like the other proform in Italian coreferring to predicates, *ci*, denotes the propositional content expressed by the causative sentence (or the sentence itself), rather than the event denoted by it. The presence of a tacit intentional predicate renders the proposition it introduces in the logical form available as the only target of reference, rather than the reference being the whole event.

## 1. *Quasi and intentionality*

### 1.1 *Ambiguity with quasi*

It has been observed that the English adverb *almost* creates different readings for a sentence in which it is inserted (McCawley 1973; cf. also Dowty 1979). The Italian counterpart of *almost*, *quasi*, has a similar behavior. Sentence (1), for example, has at least two independent readings when the subject is an intentional subject:

- (1) *Gianna ha quasi rotto il vaso.*  
 Gianna has almost broken the vase  
 “Gianna almost broke the vase.”<sup>1</sup>

The first reading is one where Gianna was about to do something that would have broken the vase, but she did not do so. In the second reading, Gianna did something that almost broke the vase, but the vase did not break. The two different readings are not typically available, however, when the subject is an unintentional, inanimate subject (with no difference if the causer is an object, like a rock, or an event, like the wind), as the sentences in (2)<sup>2</sup> illustrate:

- (2) a. *La pietra ha quasi rotto il vaso.*  
 “The rock almost broke the vase.”  
 b. *Il vento ha quasi rotto il vaso.*  
 “The wind almost broke the vase.”

The typically available reading for these sentences is one in which the rock or the wind did something that almost broke the vase, but the vase did not break. (Observe that ‘do’ is broadly used here, as in the sentence “Look what the

<sup>1</sup> In the northern variety of Italian, the complex past consistently substitutes for the simple past. The reader should not therefore be concerned with any issue related to perfectivity.

<sup>2</sup> From now on I will dispense with word-by-word glosses, unless necessary, since in this chapter I am mainly concerned with the semantics of sentences that display the same syntactic structure.

rock/wind did!,” in which no agency is implied, but rather only causation.) The exclusive reading where the rock or the wind was about to do something that would have broken the vase, but did not do it, is not available. Interestingly, when the subject is an animate unintentional causer, as in (3), the exclusive reading where Gianna was about to do something that almost broke the vase, but she did not, is not available:

- (3) *Gianna accidentalmente ha quasi rotto il vaso.*  
 “Gianna accidentally almost broke the vase.”

An animate unintentional causer thus parts with an inanimate causer with regard to the availability of the readings in question.<sup>3</sup>

1.1.1 *On empirical intuitions and vagueness.* Jean-Roger Vergnaud (personal communication, 5 May 2003) suggested that the first reading is also available whenever the subject is an unintentional causer. For example, imagine a ball that, with a given frequency, puts out spikes, which could break the vase by contact (when the ball is smooth it cannot break the vase). If the ball touched the vase right before the prickles were out, sentence (2a) (with ‘the ball’ substituting for ‘the rock’) could mean that the ball was about to do something that would have broken the vase, but it did not. But also without taking into account a spiked ball, the ‘about’ periphrasis is available. When the rock flew really close to the vase without touching it, it was about to do something, namely, to make contact with the vase, that would have broken the vase, but it did not do it. The question thus is whether a difference in logical form exists between intentional and unintentional causation. I believe it does.

First, intentional causation instantiates the occurrence of an event (an intention) sufficient to license modification by *quasi* rendered with the ‘about’ periphrasis. Second, this event is logically prior to the event describable by the ‘do’ periphrasis, that is, an intention of performing some act is an event that logically (but not *de facto*) implies the performance of an act describable by the same content of that very same intention (an intention of breaking the vase

---

<sup>3</sup> McCawley (1973) correctly thought that (1) is three-ways ambiguous, and (2) and (3) two-ways ambiguous. The third reading is obtained by breaking the situation corresponding to the second reading into two more detailed situations. As, however, this reading is available whatever the volitionality of the subject, it is not relevant in this work. The number of readings actually available depends also on the aspectual properties of the predicate, without regard to intentionality (Dowty 1979). Thus an accomplishment is two-ways ambiguous if unintentional, three-ways ambiguous if intentional. Achievement and activity predicates, on the other hand, are not ambiguous when unintentional, and two-ways ambiguous when intentional. See the discussion of sentence (4) regarding this issue.

logically implies an act of breaking the vase). If Gianna had the intention of breaking the vase without acting according to that intention (e.g., she changed her mind), sentence (1) is fine. It is not necessary for Gianna to do anything at all. The speaker/hearer of (1) simply needs to know that Gianna had the intention of breaking the vase. But Gianna might have acted on the basis of her intention, and in such a case we have the ‘do’ periphrasis. A third and final point: An intentional event is exclusively described by the ‘about’ periphrasis, that is, if Gianna just had an intention without acting on its basis, it does not seem correct to claim that she did something.<sup>4</sup>

In the case of the spiked ball, on the other hand, the imminent mutation of the ball did not actually occur and is not sufficient to license modification by *quasi*. In order for the ball to have almost broken the vase with its spikes, it is not sufficient that the ball was about to eject the spikes. Some other event that can be described with the ‘do’ periphrasis needed to occur, for example, the ball touched the vase. Moreover, the imminent mutation of the ball is not logically followed by an event describable with the ‘do’ periphrasis, since it does not imply the occurrence of any other event describable by the very same content that describes it (the event of the vase breaking can follow, but not logically, the mutation of the ball, since these two events cannot be described by the same content). To summarize, within intentional causation the event describable with the ‘about’ periphrasis actually happened, is sufficient, and is logically followed by an event describable with the ‘do’ periphrasis. With unintentional causation, instead, the event describable with the ‘about’ periphrasis did not actually happen, and is not sufficient, since an event describable with the ‘do’ periphrasis, which does not logically follow it, needs to occur. The intuition that the ‘about’ periphrasis is available also for unintentional causation thus does not capture the fact that the event described by this periphrasis did not occur, is not sufficient to license the use of *quasi*, and is not logically followed by any other event.

---

<sup>4</sup> An anonymous reviewer suggests evaluating sentence (3) against the following situation: A hidden button, if pressed, causes the vase to break and Gianna is not aware of its effect. Gianna is on the verge of accidentally pressing the button, but does not. This event is correctly described by sentence (3) with the ‘about’ periphrasis (“about to break the vase”). However, this very same event can also be described with the ‘do’ periphrasis. If we can say that Gianna was about to accidentally press the button, Gianna must have done something, for example, move her hand; therefore this situation can also be described as “Gianna did something that almost broke the vase.” For this reason, an intention is different from an accidental act: It is not an act at all, and it can exclusively be rendered with the ‘about’ periphrasis. Furthermore, the accidental pushing of the button does not logically imply the breaking of the vase, since these two events cannot be described by the same content.

Regarding the simpler event of the flying rock, it also can be described with the ‘about’ periphrasis, as previously observed. In such a situation, however, we do not have two different events describable with two different periphrases, but a single event, the rock flying close to the vase, which can be described in different ways (as almost breaking the vase, or as almost making contact with the vase). Even more clearly in this case of vagueness, we do not have an event exclusively describable by the ‘about’ periphrasis sufficient to license the use of *quasi*, and logically followed by an event describable by the ‘do’ periphrasis.<sup>5</sup>

To remain in the domain of vagueness, Dowty criticizes McCawley’s (1973) view on the ambiguity of sentences with *almost* on the basis that it is not sufficient to imagine a number of distinct situations to which the different reading of (1) could be applied. He follows Zwicky and Sadock’s (1975) argument that basing a claim for ambiguity solely on this intuition would be like “supposing that, for example, *John has a shirt*, is ambiguous between a reading ‘John has a red shirt,’ and a reading ‘John has a blue shirt’” (1979:243). These considerations, however, do not take into account the reason why a sentence like (4) does not have multiple readings, as Higginbotham (1989) notices:

(4) John almost fell.

It is possible to imagine a variety of situations in which John almost fell. For example, he tottered, slipped, and fainted, a friend rescued him, and so on. Sentence (4) differs from (1) to the extent that there is only one event (what John did that was almost a falling) and it is a matter of vagueness which event it is exactly. Sentence (1) denotes a complex event formed of various subevents. There are two interesting facts that confirm this subdivision in events on empirical grounds. First, (4) has one more reading if John had the intention of falling (e.g., as a tactic on the football field), but he refrained. In such a case the ‘about’ reading becomes available and (4) becomes ambiguous between the reading where John was about to fall on purpose, but he suddenly changed his mind, and the reading where he did something, like tottering, according to his intention of falling, but he did not succeed (a teammate

---

<sup>5</sup> Observe that the situation of the rock flying really close to the vase and missing it can be described with the ‘about’ periphrasis, but also with the ‘do’ periphrasis (the rock did something, i.e., flying really close to the vase, that almost broke the vase). In this sense the ‘about’ periphrasis is not exclusive, as discussed in fn. 4.

rescued him before he reached the ground).<sup>6</sup> Second, it is possible to overtly realize a double occurrence of *quasi*, as illustrated in (5):

- (5) *Gianna ha quasi quasi rotto il vaso.*  
 “Gianna almost almost broke the vase.”

Besides the nonrelevant reading where *quasi quasi* is an intensifier (Gianna was extremely close to breaking the vase), (5) can describe the following situation: Gianna had the intention of breaking the vase but right before the moment she started acting this plan out, she accidentally hit the vase (a sudden and violent sneeze made her lose control) and almost broke it. The first *quasi* thus modifies Gianna’s intention, while the second modifies what she did. If (5) denotes a simple event, it is not clear why it can describe this situation.

Finally, the solution to the multiple readings of (1) that Dowty (1979), following Sadock (1979), proposes, is not sufficient to capture a broad set of data. To claim, in fact, that “x almost VPs” means simply, “there is a possible world very similar to the actual world in which x VPs is true,” does not provide an explanation for the fact that some actual situations for which very similar possible worlds exist where “x VPs is true” cannot be described with sentences containing ‘almost,’ as shown in the next subsection.

1.1.2 *There must exist an event.* The use of *quasi* (but what I am claiming here is valid for *almost* as well) is licensed exclusively by the occurrence of an event involving the sentence’s arguments. The sole existence of a situation where the event denoted by the VP might have happened is not sufficient. Thus, it would be improper to use sentence (4), or its Italian translation (6), for the situation where a banana skin thrown very close to John would have made John fall, if John had stepped on it. Something must have happened to John (e.g., he tottered):

- (6) *John e’ quasi caduto.*  
 John is almost fallen  
 “John almost fell.”

It would be improper to use sentence (2) as well to describe a situation where a rock is on the edge of the roof, very likely to fall on the vase underneath, unless, for example, the rock fell close to the vase. Finally, imagine the

---

<sup>6</sup> ‘Fall’ is an achievement predicate, which becomes ambiguous when interpreted as intentional. Moreover, sentences containing activity predicates are ambiguous when intentional. (See also fn. 3.)

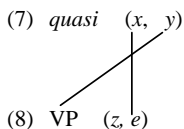
situation where Gianna is a compulsive breaker during one of her crises. It is very likely she is going to break the vase, but she does not. It would be improper to utter (1), unless we know that she intended to break the vase and she refrained. On the basis of the data in this and the previous subsection I therefore conclude that a simple modal account is not sufficient, that the existence of events is a requirement for licensing the use of *quasi*, and that the existence of different readings for the present sample of sentences is indeed a matter of ambiguity resulting from the number of subevents constituting the event complex denoted by the predicate.

The data on the ambiguity of *quasi* considered in this section will be accounted for by considering the meaning of the modifier (1.2), the logical forms of causation sentences with intentional subject and with unintentional subject (1.3), and the compositional semantics of *quasi* when in construction with causative predicates (1.4).

### 1.2 The meaning of quasi

I will develop and slightly modify a proposal made by Higginbotham (1989) for an analysis of the meaning of *almost*, a modification that takes into consideration the necessity of an event as shown in the previous subsection. *Quasi*, like *almost*, is a binary predicate taking two arguments ( $x$ ,  $y$ ) respectively true of an object and an intensional entity (either property or proposition). Its meaning can be paraphrased as “ $x$  is a thing close to (having) the property or (verifying) the proposition  $y$ .”

When *quasi*, or *almost*, modifies a causative predicate, as in the sentences I have considered, the first argument  $x$  of *quasi* is identified with an event  $e$ , and the second argument  $y$  of *quasi* is identified with the property or proposition described by the semantic content expressed by the predicate. The semantic composition is theta identification of arguments (Higginbotham 1985, 1989:481). The argument structures of *quasi* and the VP are given in (7) and (8) respectively, and the lines among them are intended to illustrate the theta identification operations:



In (9a) semantic composition as illustrated in (7) and (8) is a product of theta identification of the event  $e$  of the VP with the first argument  $x$  of *quasi*, and of the property  $\lambda e'$  VP ( $c$ ,  $e'$ ) ( $c$  being a constant) described by the semantic

content expressed by the VP with the second argument  $y$  of *almost* (syntactically, the AdvP projected by *quasi* is in V'). Thus, if the VP is, for example, 'fell,' the property  $\hat{\lambda}e'$  VP ( $c, e'$ ) is the property of falling. The form in (9a) thus represents the case in which the argument  $y$  of *quasi* is a property and is read as (9b):

- (9) a. almost ( $e, \hat{\lambda}e'$  VP ( $c, e'$ ))  
 b. Some event  $e$  is close to being an event described by the VP.

When the event  $e$  is identified with  $x$ , as in the previous case, but the proposition  $\hat{\lambda}e'$  (VP ( $c, e'$ )) is instead identified with  $y$ , semantic composition of the relevant arguments in (7) and (8) results in the form in (10a) (syntactically, the AdvP is in Infl', where the event argument  $e$  of the VP undergoes existential closure). It represents the case in which the argument  $y$  of *quasi* is a proposition and is paraphrased as (10b):

- (10) a. almost ( $e, \hat{\lambda}\exists e'$  VP ( $c, e'$ ))  
 b. Some event  $e$  is close to verifying the proposition given by the IP.

In this subsection I have illustrated the meaning of *almost* as a binary predicate true of an event and an intensional entity, either property or proposition. We will see how this account explains the ambiguity (or lack of it) of the targeted sentences once I consider their logical forms and their composition with the meaning of *quasi* in the following subsections.

### 1.3 The logical forms for the different readings

When the logical form for sentences (1) through (3) contains only one event, as with the case of the sentences with an unintentional subject, only one reading is possible. When, on the other hand, the logical form contains an event complex formed of two subevents, as in the case of the sentence with an intentional subject, two readings are available. The event complex relevant in this discussion is an ordered pair ( $e, e'$ ) (Higginbotham 2000) formed of an event  $e$  (an intention) and its direct act  $e'$ . By 'direct act' I mean what usually goes under the name of 'direct causation,' that is, the act performed by the very same person who has the intention of VP, thus not an act performed by some other person according to the subject's intention. In sentence (11a), for example, the subevent  $e$  is Gianna's intention of breaking the vase and the subevent  $e'$  is the physical act Gianna performed to break the vase. The subevent  $e'$  cannot be the act performed by someone whom, for example, Gianna ordered to break the vase:

(11) a. *Gianna ha rotto il vaso.*

“Gianna broke the vase.”

- b.  $\exists e \exists e' \{ \text{intend}(\text{Gianna}, \wedge \exists e'' \text{break}(\text{PRO}, \text{the vase}, e''), e) \& [\text{break}(\text{Gianna}, \text{the vase}, (e, e'))] \}$ .

The logical form for (11a), represented in (11b), contains the predicate ‘intend,’ denoting a relation between an individual, a proposition, and an event. In (11b) the event complex  $(e, e')$  is given by Gianna’s intention of breaking the vase and her act of breaking the vase, and the predicate ‘intend’ is a relation between Gianna, an individual, the proposition  $\wedge \exists e''$ : break (PRO, the vase,  $e''$ ) of Gianna breaking the vase, and the event  $e$ , Gianna’s intention. The logical form for sentence (12a), given in (12b), contains neither the event complex nor the predicate ‘intend’:

(12) a. *La pietra ha rotto il vaso.*

“The rock broke the vase.”

- b.  $\exists e [\text{break}(\text{the rock}, \text{the vase}, e)]$

When Gianna acts unintentionally, the logical form for sentence (11a) is similar to the one given for (12a), since in such a situation no event exists that is Gianna’s intention.<sup>7</sup>

I turn now to the logical form for the different readings for sentences (1) through (3). I first introduce the logical form for the only possible reading for sentence (2), which describes the situation where the rock did something that almost broke the vase, but the vase did not break. What I claim for sentence (2) extends to sentence (3) as well, given that both contain an unintentional subject. This form is given in (13a) and paraphrased as (13b):

(13) a.  $\exists e \{ \text{almost} [e, \wedge \lambda e' \text{break}(\text{the rock}, \text{the vase}, e')] \}$ .

- b. There is an event  $e$  that is close to being an event of the rock breaking the vase.

In (13a) there exists only one event, argument of the VP and first argument of *almost*, which takes as its second argument the property of the rock breaking the vase. The form for sentence (1) with the reading where Gianna did

---

<sup>7</sup> In fn. 3 I claimed there is one more reading for sentences (1)-(3). This complexity would require a further complication also in logical forms, since the two-reading ambiguity would require the existence of an event complex formed of an ordered pair of subevents, while the three-reading ambiguity would require the existence of an ordered triplet of subevents. As I previously said, this complexity is tangential to the issue under discussion, and I therefore do not consider it here.



something that almost broke the vase, but the vase did not break, is given in (14a), paraphrased as in (14b):

- (14) a.  $\exists e \exists e' \{ \text{intend (Gianna, } \wedge \exists e'' \text{ break (PRO, the vase, } e''), e) \ \& \ \text{almost } [(e, e'), \wedge \lambda e'' \lambda e''' \text{ (break (Gianna, the vase, } (e'', e''')))] \}$   
 b. There are two events  $e$  and  $e'$ ,  $e$  is Gianna's intention of her breaking the vase,  $e'$  is Gianna's act based on her intention, and the event complex  $(e, e')$  is close to being an event of Gianna breaking the vase.

In (14a) there exist two events forming the pair  $(e, e')$ , argument of the VP and first argument of *almost*, which takes as its second argument the property  $\wedge \lambda e'' \lambda e'''$  (break (Gianna, the vase,  $(e'', e''')$ ) of Gianna breaking the vase. Finally, I consider the logical form for sentence (1) with the reading where Gianna was about to do something that almost broke the vase, but she did not. This reading corresponds to what is usually called the 'sentential scope' of the sentence. The form is given in (15a) and (15b) is its English paraphrase:

- (15) a.  $\exists e \{ \text{intend (Gianna, } \wedge \exists e'' \text{ break (PRO, the vase, } e''), e) \ \& \ \text{almost } [e, \wedge \lambda e'' \exists e' \text{ (break (Gianna, the vase, } (e'', e')))] \}$   
 b. There is an event  $e$  that is Gianna's intention of breaking the vase and this event  $e$  is close to verifying the proposition that there is an event of Gianna breaking the vase.

In (15a) there exists only one event, the intention  $e$ , as the wide scope of the existential operator quantifying over it shows. The existential operator quantifying over the event  $e'$  corresponding to Gianna's act is in the scope of the intension operator, which creates the proposition that there is an event of Gianna breaking the vase.

As observed in section 1.1.2, the existence of an event is essential to license modification with *quasi*. This requirement is represented in the logical forms of the possible readings for sentences (1)-(3), in each one of which the existential quantification over at least one event has wide scope. The sentential scope reading for sentence (2) is not, as we saw, available. The reason for the missing reading becomes evident as soon as we consider the logical form it would have, were it possible, in (16):

- (16)  $\exists? \text{ almost } [?, \wedge \exists e \text{ break (the rock, the vase, } e)]$

The question marks in (16) mean to capture the lack of an event to existentially quantify over, and to identify with the first argument of *quasi*, which, according to the meaning of *quasi*, needs to be an event. The existential quantification over the only available event argument  $e$  in (16) needs to be in

the scope of the intension operator to deliver the sentential scope. There is no event complex ( $e, e'$ ), in other words, where  $e$  can be existentially closed and identified with the argument  $x$  of *quasi* and the existence of  $e'$  is in the scope of the intension operator to constitute the proposition identified with argument  $y$  of *quasi*. The meaning of *quasi* is therefore not satisfied.

On the basis of ambiguity phenomena with *quasi* I have claimed that there exists a tacit intentional predicate. We would expect that this covert predicate would manifest in other areas of the language. This is indeed the case. One of these areas is constituted by some phenomena with the Italian periphrastic causative *fare*, which I will describe in the next section. These phenomena appear only when the subject of *fare* is an intentional agent, and do not occur when the subject is an unintentional causer. The second area I will consider is the discourse reference of the proform *lo* in construction with *fare* (*farlo* “do so”), the use of which becomes fully acceptable only when the intention of the subject of the matrix clause is contextually made salient.

## 2. *Agent versus causer with the Italian periphrastic causative fare*

An intentional subject of *fare* generates special conditions on causation, conditions that do not exist when the subject is unintentional. These conditions concern the force dynamics of the situations described by the sentences containing the periphrastic construction. I claim that what triggers these conditions is the presence of the tacit intentional predicate introduced in section 1 in combination with the contrastive nature of *fare* with respect to the lexical causative. In the following sentences I exemplify the *fare* construction with an intentional agent (17), with a causer (18), and with an unintentional agent (19):

- (17) *Gianna ha fatto rompere il vaso.*  
Gianna has made break the vase  
“Gianna made the vase break.”
- (18) *La pietra/il vento ha fatto rompere il vaso.*  
the rock the wind has made break the vase  
“The rock/the wind made the vase break.”
- (19) *Gianna ha fatto accidentalmente rompere il vaso.*  
Gianna has made accidentally break the vase  
“Gianna accidentally made the vase break.”

The relevant interpretation of these sentences is one where the subject of *fare* breaks the vase herself/itself.<sup>8</sup> Gianna or the rock, for example, may have broken the vase by pushing it off the edge of the table, or by magic force. There is a particular type of situation, however, that can be described by sentences (18) and (19), but not by sentence (17). This type of situation is one where the subject of *fare*, for example, broke the vase by hitting it in a continuous mechanical way. Continuous mechanical causation occurs whenever the agent mechanically causes the event denoted by the embedded VP by contact at time  $t$ , and the vase breaks at time  $t$ .<sup>9</sup> Notice that none of the causal dynamics that are possible when the subject of *fare* is intentional conform to this definition of continuous mechanical causation. When Gianna pushes the vase off the edge of the table to break it, the pushing occurs at time  $t$ , and the vase breaks at time  $t'$  as a direct result of the vase falling. In the case of the use of magic, there is no mechanical causation involved. Hence, it seems that the ban on continuous mechanical causation emerges only when *fare* has an intentional subject. This phenomenon is highly widespread, being characteristic of the *fare* construction when it embeds an unaccusative, or alternating unaccusative, predicate.<sup>10</sup>

---

<sup>8</sup> There is another interpretation where the intentional subject of *fare* had someone else break the vase. As this interpretation is not relevant for this chapter, I will not consider it further. See Vecchiato (2003) for a discussion of this interpretation.

<sup>9</sup> Sentence (17) needs to be considered against situations carefully designed according to the definition of continuous mechanical causation as given here. If any of the elements in the definition is not represented in a situation created for testing (17), the sentence becomes available. Thus, an anonymous reviewer, on the basis of languages closely related to Italian, claims that (17) can be used to refer to a situation where the repeated improper use of the vase finally causes it to break. This description is unfortunately too vague for me to imagine a concrete situation. Even if, however, the situation the reviewer has in mind, whatever it is, is one of intentional mechanical causation, it is not a case of 'continuous' mechanical causation as I define it, if we want to take seriously the reviewer's phrase "repeated improper use." Continuous mechanical causation occurs, as stated earlier, "whenever the agent mechanically causes the event denoted by the embedded VP by contact at time  $t$ , and the vase breaks at time  $t$ ." In other words, the definition of continuous mechanical causation for the VP 'break the vase' involves only one act of breaking the vase, co-occurring with the result of the vase breaking. If it is Gianna's repeated improper use that breaks the vase, there are a number of events when Gianna improperly used the vase, each contributing to the breaking of the vase. Suppose the breaking of the vase occurs at time  $t$ . There are various acts of improper use occurring at times  $t_1, t_2, \dots, t_n$ , before  $t$ , which contribute to the breaking. This situation, therefore, does not conform to the definition of continuous mechanical causation.

<sup>10</sup> See Vecchiato (2003) for an extensive description of force dynamic constraints with the Italian periphrastic causative and its plausibility with respect to other notions traditionally used to account for the data. See Levin (1993) for a list of unaccusative and alternating predicates in English (Italian has almost, but not exactly, the same list).

I believe the tacit intentional predicate revealed by the ambiguity with *quasi*, in combination with the contrastive nature of the periphrastic causative, renders causal dynamics and the discrimination among different types of them relevant, somewhat similarly to what happens when an overt intentional predicate is present. It is in fact characteristic of an intentional predicate to manifest causal dynamics that are otherwise not emerging.

Typically, sentence (20) entails (21), since a peanut butter and jelly sandwich is made of peanut butter, jelly, and bread:

- (20) Gianna ate the peanut butter and jelly sandwich.
- (21) Gianna ate the peanut butter, the jelly, and the bread.

If, however, the predicate ‘eat’ is embedded under an overt intentional predicate, the entailment is lost, as sentences (22) and (23) show:

- (22) Gianna tried to eat the peanut butter and jelly sandwich.
- (23) Gianna tried to eat the peanut butter, the jelly, and the bread.

Sentence (22) does not entail sentence (23), since they can be taken to report different intentions and therefore expectations on how Gianna will behave, how she imagines the task before her, and how she divides it up into events. The intention reported by (22), for example, might be eating the sandwich as a whole; (23), instead, could be taken to report Gianna’s intention of eating the peanut butter separately from the jelly, and finally the bread. Now, if we reconsider sentences (20) and (21) and focus on Gianna’s intentions, the tendency is to describe Gianna’s action in a way faithful to what she tried to do, as with sentences (22) and (23), with the difference that Gianna not only tried to perform the tasks expressed by the predicates, but also succeeded in doing so.

To summarize this section, the presence of a tacit intentional predicate in combination with the contrastive nature of the Italian periphrastic makes a fine-grained causal dynamics of events relevant, in a guise similar to what an overt intentional predicate or focus on the subject’s intention will do. In the next section we will consider another phenomenon in Italian where the focus on the subject’s intentions is crucial.

### 3. *The proform lo + fare*

The Italian proform *lo*, when coreferring with a predicate rather than with an object, seems to refer to the propositional content expressed by an IP or CP, rather than to the event described by the sentence. I illustrate this issue by investigating the use of this proform in combination with *fare*. The proform *lo*

in the second conjunct in sentence (24) corefers with the VP *rompere il vaso* in the first conjunct:<sup>11</sup>

- (24) ??*Gianna ha* [<sub>VP</sub> *rotto la finestra*]<sub>i</sub>, *ma Maria non lo<sub>i</sub> ha fatto.*  
 Gianna has broken the window but Maria not it has done  
 “Gianna broke the window, but Maria did not do so.”

Speakers’ grammaticality judgment<sup>12</sup> of (24) is typically that they understand the content, but the sentence is not completely acceptable, and they would rather use (25) instead, where the ellipsis in the second conjunct expresses the content of breaking the window:

- (25) *Gianna ha rotto la finestra, ma Maria no.*  
 Gianna has broken the window but Maria no  
 “Gianna broke the window, but not Maria.”

If, however, the situation emphasizes the subjects’ intentions, (24) becomes not only acceptable, but a more appropriate and specific way of describing the situation than (25). The situation I proposed to the speakers is one where Gianna and Maria belong to a gang, which has the intention of breaking a window in a house. Gianna proceeded to do so, but Maria did not act according to the plan. In this situation the women’s plan is contextually focused, and because it is the only element added to the situation I originally asked speakers to consider (one where Gianna broke the window and Maria did not without mention of their intentions), it is what renders the sentence acceptable. It is thus plausible that the proform *lo* is licensed by the tacit intentional predicate presented in the previous two sections, and does not refer to the whole event, formed of intention, act, and result (see (11b) for the logical form of the first conjunct in (24)).<sup>13</sup>

<sup>11</sup> The Italian verb *fare* can be translated, according to the context in which it appears, both as English “make” (see section 2 for this use) and English “do,” as in (24).

<sup>12</sup> These grammaticality judgments were collected from six native speakers, including myself. They were also confirmed by the Italian audience on various occasions in which I presented these findings.

<sup>13</sup> Consider that intentionality is not entailed by the light verb *fare* itself, but rather, as we will see next, it is a condition for the felicitous reference of the proform *lo*. If intentionality were entailed by *fare*, we would have two lexical entries for *fare*: one for *fare* in constructions with *lo*, which would have the feature ‘intentional’, and one for *fare* in causative constructions, which, as seen in section 2, does not have this feature, as it can occur with unintentional subjects. This seems to be an unnecessary ad hoc complication in the lexicon. See also fn. 14 for an empirical argument against this position.

When asked to judge the original situation, the speakers observed that the appropriate way to convey it is by means of (25). The elliptical content in the second conjunct must refer, thus, to the whole event expressed by the predicate, whether intentional (26) or not (27):

- (26)  $\exists e \exists e'$  {intend (Gianna,  $\hat{\exists}e$  "break (PRO, the window,  $e$ "),  $e$ ) & [break (Gianna, the window, ( $e$ ,  $e'$ ))]} &  $\neg \exists e'' \exists e'''$  {intend (Maria,  $\hat{\exists}e$  "break (PRO, the window,  $e$ "),  $e$ ) & [break (Maria, the window, ( $e''$ ,  $e'''$ ))]}  
 (27)  $\exists e$  [break (Gianna, the vase,  $e$ )] &  $\neg \exists e'$  [break (Maria, the vase,  $e'$ )]

We can, in fact, use (25) whether the breaking is intentional or unintentional. This possibility is a priori ruled out with (24), since the focus on the intention is what makes this sentence acceptable. In fact, there is no situation that can improve sentence (28):

- (28) ?\**Gianna ha accidentalmente [rotto la finestra]<sub>i</sub>, ma Maria non lo<sub>i</sub> ha fatto.*  
 "Gianna accidentally broke the window, but Maria did not do so."

The same applies to similar sentences with a nonanimate causer:

- (29) ?\**La pallottola ha [rotto la finestra]<sub>i</sub>, ma la pietra non lo<sub>i</sub> ha fatto.*  
 "The bullet broke the window, but the rock did not do so."

Interestingly, the only existing Italian proforms that can corefer with predicates, *lo* and *ci*, need to be licensed by predicates embedding opaque contexts or propositions. A context is opaque when the form of the embedded sentence or the content of the proposition becomes semantically relevant, rather than the event described by it. Sentences (30) and (31) contain predicates taking opaque contexts or propositions:

- (30) *Gianna ha detto/suggestito di [PRO<sub>arb</sub> scrivere un libro]<sub>i</sub>, e anche Maria lo<sub>i</sub> ha detto/suggestito.*  
 "Gianna said/suggested to write/writing a book, and also Maria said/suggested it."  
 (31) *Gianna ha provato a/pensato di [PRO scrivere un libro]<sub>i</sub>, e anche Maria ci<sub>i</sub> ha provato/pensato.*  
 "Gianna tried to/thought about write/writing a book, and also Maria tried to/thought about it."

Predicates taking a transparent context, that is, a context where the form of the embedded sentence and the content of the proposition do not have compositional semantic import (even if they, of course, have cognitive relevance), do not license either *lo* or *ci*, as sentence (32) illustrates:

- (32) \**Gianna ha iniziato a/smesso di [PRO scrivere un libro]<sub>i</sub>, e anche Maria ci/lo<sub>i</sub> ha iniziato/smesso.*  
 “Gianna started/finished writing a book, and also Maria started/finished.”

It is then plausible that a predicative proform does not refer to an event, but rather to the embedded sentence or proposition itself.<sup>14</sup> Sentences (28) and (29), and sentence (24) without considering the appropriate situation, are not completely acceptable because there is nothing that separates the sentence or proposition from the event it denotes to be the reference of *lo*. Speakers refer to the whole event, as in (26) and (27). When, on the other hand, the appropriate situation underlines the tacit intentional predicate in (24), the proposition, which is one of the terms in the relation, becomes visible for reference. The logical form in (33) claims that Gianna intentionally broke the window but Maria did not act according to the proposition expressing the event of breaking the window:

- (33)  $\exists e \exists e' \{ \text{intend}(\text{Gianna}, \wedge \exists e'' \text{ break}(\text{PRO}, \text{the window}, e''), e) \ \& \ [\text{break}(\text{Gianna}, \text{the window}, (e, e'))] \} \ \& \ \neg \text{fare}(\text{Maria}, \wedge \exists e'' : \text{break}(\text{PRO}, \text{the window}, e''))$

It is important to realize that the proform *lo* refers to the embedded sentence or proposition it expresses, which is a term in the relation constituted by the predicate ‘intend.’ It does not refer to the whole predicate ‘intend.’ Sentence (34) is in fact possible, the intentional predicate in the first conjunct being explicitly underlined by the phrase *di proposito* and its contrastive use:

- (34) *Gianna ha [rotto la finestra]<sub>i</sub> di proposito, Maria invece lo<sub>i</sub> ha fatto accidentalmente.*  
 “Gianna broke the window on purpose, Maria instead did so accidentally.”

The correct logical form for (34) is (35), not (36):

- (35)  $\exists e \exists e' \{ \text{intend}(\text{Gianna}, \wedge \exists e'' \text{ break}(\text{PRO}, \text{the window}, e''), e) \ \& \ [\text{break}(\text{Gianna}, \text{the window}, (e, e'))] \} \ \& \ \text{on purpose}(e, e') \ \& \ \exists e''' [\text{fare}(\text{Maria}, \wedge \exists e'' : \text{break}(\text{PRO}, \text{the window}, e''), e''') \ \& \ \text{accidentally}(e''')]$

<sup>14</sup> When *fare + lo* is embedded under a modal predicate, it can co-occur with a nonagentive causer:

(i) *La pallottola puo' [PRO rompere la finestra]<sub>i</sub>, ma la pietra non puo' farlo<sub>i</sub>*  
 “The bullet can break the window, but the sun cannot do so.”

A predicate embedded under a modal does not refer to an actual event. It refers to an event in possible worlds defined by the content of the sentence.

- (36)  $\exists e \exists e'$  {intend (Gianna,  $\wedge \exists e''$  break (PRO, the window,  $e''$ ),  $e$ ) & [break (Gianna, the window, ( $e, e'$ ))] & on purpose ( $e, e'$ )} &  $\exists e'''$  [fare (Maria, intend (Maria,  $\wedge \exists e''$  break (PRO, the window,  $e''$ ),  $e'''$ )) & accidentally ( $e'''$ )]

The logical form in (35), in fact, states that there was a complex intentional event of Gianna breaking the window, that this event was on purpose, that Maria performed an event described by the content of the proposition  $\wedge \exists e''$ : break (PRO, the window,  $e''$ ), and that this event was accidental. The logical form in (36), on the other hand, is a contradiction, as in its third conjunct it states that Maria performed an intentional event of the type described by the content of the proposition  $\wedge \exists e''$ : break (PRO, the window,  $e''$ ) and this event was accidental.<sup>15</sup>

#### 4. *Conclusions*

The existence of a tacit intentional predicate accounts for the ambiguity of sentences containing *quasi*, the ban on some particular causal dynamics with causative *fare*, and the interpretation of the predicate proform *lo* in Italian, all phenomena for which the asymmetry between intentional and unintentional causation is crucial. These phenomena are language-internal arguments for the existence in Italian of a tacit intentional predicate. The distinction between intentional and unintentional causation is, however, overt in other languages.

Various languages display different ways to distinguish between an intentional agent and an unintentional causer. Two ways have been brought to my attention. Tagalog and Malagasy, two western Malayo-Polynesian languages, discriminate between two diverse causative morphemes (Travis 2000). In Marathi, an Indo-Aryan language, an intentional agent is the subject of the sentence and takes the ergative agentive case in the past, while an unintentional causer cannot be the subject and is realized as an oblique phrase (Joshi 1993). The existence of languages that have an overt grammatical distinction between intentional and unintentional subjects suggests that the covert distinction in Italian is plausible. On the basis of these languages, where the difference of agent versus causer seems to be syntactic as well as semantic, it would be interesting to explore whether there are language-internal data for a syntactic distinction between agent and causer. The interactions of the tacit intentional predicate with negation and adverbs other than *quasi* are pertinent to this future inquiry.

---

<sup>15</sup> I thank an anonymous reviewer for bringing to my attention sentence (34), which provides support for my analysis if correctly interpreted with (35), and not with (36).



## REFERENCES

- Dowty, David R. 1979. *Word Meaning and Montague Grammar: The semantics of verbs and times in generative semantics and in Montague PTQ*. Dordrecht: Reidel.
- Higginbotham, James. 1985. "On Semantics". *Linguistic Inquiry* 16, 547-93.
- , 1989. "Elucidation of Meaning". *Linguistics and Philosophy* 12, 465-517.
- , 2000. "Accomplishments". *Proceedings of GLOW in Asia II* ed. by Marmoru Saito, 72-82. Nagoya, Japan: Nanzan University.
- Joshi, Smita. 1993. *Selection of Grammatical and Logical Functions in Marathi*. Ph.D. dissertation, Stanford University.
- Levin, Beth. 1993. *English Verb Classes and Alternations: A preliminary investigation*. Chicago: University of Chicago Press.
- McCawley, James. 1973. "Syntactic and Logical Arguments for Syntactic Structure". *Three Dimensions of Linguistic Theory* ed. by Osamu Fujimura, 259-376. Tokyo: TEC.
- Sadock, Jerrold. 1979. "Almost". *Radical Pragmatics* ed. by Peter Cole, 257-272. New York: Academic Press.
- Travis, Lisa. 2000. "Events Structure in Syntax". *Events as Grammatical Objects* ed. by Carol Tenny & James Pustejosky, 145-182. Stanford, CA: Center for the Study of Language and Information.
- Vecchiato, Antonella. 2003. "The Italian Periphrastic Causative and Force Dynamics". *USC Working Papers in Linguistics* 1, 93-109.
- , To appear. "Causation and Intentionality". *Proceedings of the 31<sup>st</sup> Western Conference on Linguistics (WECOL)* ed. by Brian Agbayani, Vida Samiiian, & Benjamin V. Tucker. Fresno: Department of Linguistics, California State University.
- Zwicky, Arnold & Jerrold Sadock. 1975. "Ambiguity Tests and How to Fail Them". *Syntax and Semantics* ed. by John P. Kimball, vol. IV, 1-36. New York: Seminar Press.

# INVERSION, RECONSTRUCTION, AND THE STRUCTURE OF RELATIVE CLAUSES\*

LUIS VICENTE  
*Leiden University*

## 0. *Introduction*

The goal of this chapter is to argue in favor of a double-headed analysis of restrictive relative clauses (Platero 1974; Sauerland 1998, 2000). One of the head positions is the SpecCP slot of the relative clause, while the other is the NP the relative clause is adjoined to. Contra Platero and Sauerland, I argue that there is only one head position, the other one being taken up by an empty, expletive-like element that I dub '*eNP*.' I also argue for Sauerland's and Aoun and Li's (2003) hypothesis that both the matching and the raising analyses are made available by UG. However, whereas they assume that each analysis involves a different syntactic structure, I argue that there is only one structure, the differences between 'raising' and 'matching' arising from an indeterminacy on the way in which the 'real' relative head and *eNP* are merged. Moreover, this indeterminacy is also shown to be responsible for certain reconstruction asymmetries in Spanish relative clauses that had not been noticed so far.

The article is organized as follows. In section 1, I introduce the relevant data set from Spanish and compare it to its English counterpart. Section 2 consists of a brief review of the double-headed analysis of relative clauses, and the modifications I introduce. In section 3, I take a brief excursus so as to link the variations in word order within relative clauses to variations in word order in other parts of the language. In section 4, I return to the asymmetries introduced in section 1, and show how they can be accounted for in terms of the theory developed so far. Finally, section 5 contains a number of open questions that I leave for future research.

---

\* I want to thank all those who provided comments and discussion on earlier incarnations of the manuscript, leading to a considerably better version of it: in reverse alphabetical order, Mark de Vos, Joanna (Ut-Seong) Sio, Johan Rooryck, Jairo Nunes, Anikó Lipták, Marjo van Koppen, Alex Grosu, Crit Cremers, Jeroen van Craenenbroeck, Lisa Cheng, the audiences at the 12<sup>th</sup> Manchester Postgraduate Conference and LSRL 33, and three anonymous reviewers. Special thanks to Josu Vitores, always ready to check Spanish judgments with me. I am also grateful to the Leiden University Fund and the Leiden Centre for Linguistics for their financial support. Any remaining mistakes, misconceptions, and shortcomings are my own.

## 1. *Some asymmetries in English and Spanish*

### 1.1 *English*

The discovery of the reconstruction/antireconstruction contrast in (1) is usually attributed to Williams and van Riemsdijk (1981). Descriptively, an R-expression contained in a complement PP, embedded in turn in a fronted *wh* phrase, is interpreted as if it were in its base position, hence the Condition C violation in (1a). On the other hand, if the R-expression belongs in a relative clause, as in (1b), no reconstruction takes place, and coreference is possible.

- (1) a.\* [Which picture [<sub>PP</sub> of John<sub>i</sub>]] did he<sub>j</sub> see *t*?  
 b. [Which picture [<sub>RC</sub> that John<sub>i</sub> took]] did he<sub>j</sub> see *t*?

Particular details aside, many analyses of reconstruction (Lebeaux 1988, 1992; Chomsky 1993; Heycock 1995; Epstein et al. 1998; Sauerland 1998, 2000; Fox 1999; Stepanov 2001a, 2001b) assume that the offending R-expression is in the c-command domain of the binder at the relevant level of representation in (1a), whereas this is not the case in (1b). Here I adopt the spirit of Lebeaux's (1988, 1992) analysis. He argues that (1) stems from a dichotomy between complements (PPs) and adjuncts (relative clauses and some PPs), ultimately reducible to theta assignment. Specifically, he argues that complements, since they receive a theta role, must be present at the level where theta relations are expressed, namely, D-Structure. On the other hand, adjuncts do not receive theta roles, therefore their insertion can be delayed until S-Structure. Stepanov (2001a, 2001b) reinterprets Lebeaux's proposal in a DS/SS-less system as cyclic versus postcyclic insertion: Cyclically inserted phrases enter the derivation before their heads undergo further operations; postcyclically inserted ones adjoin to their heads at a later point.<sup>1</sup> Under this hypothesis, together with the assumption that reconstruction is LF activation of a lower, unpronounced copy (Chomsky 1993; Bobaljik 2002), we arrive at the (simplified) LF representations in (2), which derive the asymmetries in (1).

- (2) a. [which *x*] he saw [*x* picture [of John]]  
 b. [which *x* [that John took]] he saw [*x* picture]

The contrast between PPs and relative clauses follows from the assumption that binding theory applies at LF (Chomsky 1993; Fox 1999). In (2a), *John* is c-commanded by a coreferential pronoun, therefore causing a Condition C

---

<sup>1</sup> Stepanov (2001a, 2001b) argues that adjuncts must be inserted postcyclically. However, my analysis will rely on weakening *must* to *may*. This option has also been explored (e.g.) by Lebeaux (1992) and Fox (1999).

violation. In (2b), on the other hand, there is no copy of the relative clause containing *John* in the c-command domain of *he*, so binding theory is satisfied. This is the analysis I will follow for the remainder of this chapter.

### 1.2 Spanish

As shown in (3), Spanish PPs show reconstruction effects in the same way as their English counterparts. Therefore, I will ignore them in this chapter. It is the relative clauses in (4) that show a more intriguing pattern: Anti-reconstruction effects parallel to (1b) arise only as long as the SV order is maintained inside the relative clause (4a). If the order is VS, they reconstruct (4b). Example (4c) is provided to show that swapping the positions of the name and the pronoun renders the sentence grammatical. I take this to indicate that the ungrammaticality of (4b) is, effectively, due to reconstruction, and not something else.<sup>2</sup>

- (3) \*¿[ *Qué foto* [ *de Juan<sub>i</sub>*]] *ha visto él<sub>i</sub> t?*  
 what picture of Juan has seen he
- (4) a. ¿[ *Qué libro* [ *que Juan<sub>i</sub> escribió*]] *ha publicado él<sub>i</sub>?*  
 what book that Juan wrote has published he
- b. \*¿[ *Qué libro* [ *que escribió Juan<sub>i</sub>*]] *ha publicado él<sub>i</sub>?*  
 what book that wrote Juan has published he
- c. ¿[ *Qué libro* [ *que escribió él<sub>i</sub>*]] *ha publicado Juan<sub>i</sub>?*  
 what book that wrote he has published Juan

The sentences in (4) constitute the empirical contribution of this chapter, since they have not been discussed anywhere in the literature. With regard to this set of data, José Camacho (personal communication, n.d.) suggests that, since Romance postverbal subjects are usually focused, information structure (IS) considerations may force a disjoint reference in (4b). I will not consider this option. For one, (4c) indicates that we can account for this phenomenon with a purely configurational—and therefore simpler—version of binding

<sup>2</sup> Interestingly, this effect also shows up in French and Italian. I am grateful to Johan Rooryck and Denis Delfitto (personal communications, n.d.) for constructing the following examples for me:

- (i) a. [*Quale fotografia* [ *che Gianni ha fatto*]] (*pensi que lui<sub>i</sub> abbia visto?*)  
 what picture that Gianni has made think that he has seen
- b. \*[*Quale fotografia* [ *che ha fatto Gianni*]] (*pensi que lui<sub>i</sub> abbia visto?*)  
 what picture that has made Gianni think that he has seen
- (ii) a. [*Quelle histoire* [ *que Jean a racontée*]] *a-t-il<sub>i</sub> inventée de toutes pieces?*  
 what story that Jean has told has-he invented of all pieces
- b. \*[*Quelle histoire* [ *qu'a racontée Jean*]] *a-t-il<sub>i</sub> inventée de toutes pieces?*  
 what story that-has told Jean has-he invented of all pieces

theory. Moreover, in an IS approach, what is relevant is that the postverbal subject is focused. It is not clear to me whether its binding-theoretic status should be taken into account. If it is not, we would predict that (4c) should be ungrammatical in the same way as (4b), contrary to fact. If it is, we would get the right results, but we would simply be reduplicating something we can achieve independently with a strictly configurational binding theory.<sup>3</sup>

To sum up, the data from English show that there is a correlation between reconstruction and cyclic insertion on the one hand and antireconstruction and postcyclic insertion on the other. Since (4b) shows reconstruction effects, we can assume that VS relatives are inserted cyclically, as opposed to SV relatives. This chapter is an attempt to answer why reconstruction is related to inversion in this way.

## 2. *On the analysis of relative clauses*

### 2.1 *Competing approaches*

There exist in the transformational literature two major trends in the analysis of relative clauses. One of them is the ‘empty operator’ or ‘matching’ analysis, the classical reference for which is Chomsky (1977). According to this approach, the head NP is base generated in its base position, and the relative itself is a CP adjoined to the head.<sup>4</sup> The position the head should have occupied inside the relative is taken up by a phonetically unrealized  $\lambda$  operator that moves to SpecCP. This movement turns the CP into a predicate. Thus, the set the CP denotes can intersect with the set denoted by the head and yield the right interpretation, as in (5).

(5) The matching analysis:

The  $[_{NP} [_{NP} \text{girl}] [_{CP} Op_i \text{ that I saw } t_i]]$  (where  $i = j$ )

The second line of analysis is usually referred to as the ‘raising’ or ‘promotion’ approach. Its first detailed layout is Vergnaud (1974), and it has been revived in the last decade by Kayne (1994) and much subsequent work.<sup>5</sup> The core ideas of this analysis are (a) that the head NP is generated in the gap position and raises to the SpecCP position and (b) that the relative CP is a

<sup>3</sup> As I was finishing this chapter, I came across Gutierrez Bravo (2002, 2003), who gives an interesting analysis of word order in Mexican Spanish in terms of IS. He does not discuss reconstruction phenomena, however. Unfortunately, time constraints prevent me from including an evaluation of the extent to which his work is (in)compatible with my hypothesis.

<sup>4</sup> The usual analysis is that it is adjoined to NP in restrictive relatives and to DP in appositives.

<sup>5</sup> The most thorough analysis of relative clauses under this approach is probably Bianchi (1999, 2000). See also many of the papers in Alexiadou et al. (2000), de Vries (2002), and Bhatt (2002).

complement of the external determiner.<sup>6</sup> The conjunction of these two hypotheses yields the right word order in (6).

- (6) The raising analysis:  
the [<sub>CP</sub> girl<sub>i</sub> that I saw *t<sub>i</sub>*]

The implicit assumption in the literature is that one of these analyses should suffice to account for all (restrictive) relatives, and the discussion has focused on which one is more appropriate, with equally strong arguments on both sides. However, Sauerland (1998, 2000) and Aoun and Li (2003) have proposed that both analyses are actually made available by UG and are necessary to account for different properties of relativization. Nonetheless, while Sauerland and Aoun and Li argue that both the matching and the raising analyses exist as independent syntactic entities—that is, (5) and (6)—I make a stronger claim in this chapter. My proposal is that there is one single structure for relative clauses, but there are two possible ways to construct it. Each of these options yields a derivation with different properties, which can account for the differences predicted by having two separate structures.

## 2.2 *The ambiguity of English relatives*

Sauerland (1998, 2000) observes that English restrictive relatives are ambiguous between the matching and the raising analyses. On the one hand, sentences like (7) call for a matching analysis, where the head *relatives of John* is not fully represented in the gap position. Otherwise, we would expect a Condition C violation. Since the sentence is grammatical, Sauerland concludes that what is in the gap position must be an impoverished representation of the head, rather than a full copy.

- (7) The [<sub>relatives of John<sub>i</sub></sub>][that he<sub>i</sub> hates *e*]

On the other hand, data like (8) require a raising analysis, which leaves a full copy of the head in the gap position. In (8a), anaphor binding requires a copy of *each other* to be in the c-command domain of the binder. Similarly, the reading of (8b) in which *every doctor* takes scope over *two patients* requires that a copy of the latter be in the gap position.

- (8) a. The [<sub>interest in each other<sub>i</sub></sub>][that John and Mary<sub>i</sub> showed *e*]  
b. I saw the [<sub>two patients</sub>][that every doctor visited *e* today]  
(every doctor >> two patients)

---

<sup>6</sup> Though see Platzack (2000), who argues that the relative CP is a complement to an external NP.

Sauerland's (1998, 2000) solution to this paradox is reminiscent of Platero's (1974) analysis of Navajo relatives,<sup>7</sup> which can be either externally or internally headed. His proposal is that there are two instances of the head—one inside and one outside the relative CP—and one of them must undergo deletion.<sup>8</sup> If it is the CP-external head that deletes (9b), we get an internally headed relative; otherwise an externally headed one (9c).

- (9) a. [[<sub>CP</sub> *ashkii at'ééd yiyiiltsa nee*] *ashkii yalti'* (base structure)  
           boy girl saw C boy speaks  
           "The boy that the girl saw is speaking."  
    b. [[<sub>CP</sub> *ashkii at'ééd yiyiiltsa nee*] ~~*ashkii*~~ *yalti'* (internally headed)  
    c. [[<sub>CP</sub> ~~*ashkii*~~ *at'ééd yiyiiltsa nee*] *ashkii yalti'* (externally headed)

Acknowledging the existence of both matching and raising relatives in English, Sauerland (1998, 2000) adopts Platero's (1974) double-headed analysis for the former type, but with two amendments. First, he stipulates that in English it is always the internal head that undergoes deletion. That is, English matching relatives always have the structure in (9c) for Navajo, obviating the pre-/post-nominal distinction. This claim is modeled in Bresnan's (1973) analysis of English comparatives, schematized in (10b).<sup>9</sup> The difference seems to be that the deletion rule targets nouns/NPs in one case and adjectives/APs in the other.

- (10) a. I saw a girl [<sub>CP</sub> ~~girl~~ that was wearing a red dress]  
       b. The table is wider than the rug is ~~wide~~

The second difference is that Sauerland (1998, 2000) assumes a 'vehicle change' (VC) operation (see vanden Wyngaerd & Zwart 1991 and Fiengo & May 1994), whereby "an R-expression or wh- trace in the antecedent of ellipsis can correspond to a pronoun in the elided material" (Sauerland 2000:13).<sup>10</sup> The pair in (11) is given as evidence for this process.

<sup>7</sup> Actually, Sauerland (1998, 2000) does not mention Platero's (1974) analysis, but the core idea in both approaches is obviously the same. Thanks to Alex Grosu (personal communication, n.d.) for mentioning Platero's paper to me.

<sup>8</sup> Bear in mind that these copies are related not derivationally, but through a construal rule. In present-day terms, they would constitute separate items in the numeration. For this reason, I will refer to them as 'instances' or 'tokens,' rather than 'copies.'

<sup>9</sup> Though see Chomsky's (1977) analysis of comparatives, where he makes use of an empty operator, so only one adjective is present and the deletion rule is unnecessary.

<sup>10</sup> VC works on the hypothesis that *wh* traces and R-expressions are [-pronominal, -anaphoric] elements. VC changes the [pronominal] feature value, yielding a [+pronominal, -anaphoric] representation that corresponds to pronouns and *pro*.

- (11) a.\* The [story about Mary<sub>i</sub>] [that she<sub>i</sub> told *e*]  
 b. The [story about Mary<sub>i</sub>] [that she<sub>i</sub> thinks [Peter told *e*]]

One consequence of turning an R-expression into a pronoun is a change in its binding theoretic status: It goes from being subject to Condition C to abiding by Condition B. This seems to be the cause of the contrast in (11). Because of VC, *Mary* changes to *her* in the internal head, and after reconstruction to the gap position, this head is in the c-command domain of *she*. In (11a), the gap is in the same binding domain as the binder, hence the Condition B violation. However, the addition of a second embedding in (11b) creates a new domain in which Condition B cannot operate. As a consequence, the latter sentence is grammatical.<sup>11</sup>

### 2.3 More SV/VS asymmetries in Spanish

The goal of this section is to apply the previously described tests to Spanish in order to get a fuller set of data. We begin by noticing that, if the SV order is maintained, the judgments for the Spanish sentences parallel those for English (12).

- (12) a. *Los [parientes de Juan<sub>i</sub>] [que él<sub>i</sub> odia *e*] viven lejos de aquí*  
 the relatives of Juan that he hates live far from here  
 b. *El [interés del uno en el otro<sub>i</sub>] [que María y Juan<sub>i</sub> mostraron *e*]*  
 the interest of one in the other that María and Juan showed  
 c. *Vi a [los dos pacientes] [que cada médico visitó *e* hoy]*  
 saw to the two patients that each doctor visited today  
 (*cada medico >> dos pacientes*)  
 d. \**La [historia sobre María<sub>i</sub>] [que ella<sub>i</sub> contó *e*]*  
 the story about María that she told  
 e. *La [historia sobre María<sub>i</sub>] [que ella<sub>i</sub> piensa [que Pedro contó *e*]]*  
 the story about María that she thinks that Pedro told

However, the judgments are different for some of the VS counterparts of the sentences in (12), shown in (13).

- (13) a. \**Los [parientes de Juan<sub>i</sub>] [que odia él<sub>i</sub> *e*] viven lejos de aquí*  
 the relatives of Juan that hate she live far from here

<sup>11</sup> Jeroen van Craenenbroeck (personal communication, n.d.) asks why (11a) is ungrammatical but (7) is not, given that, at first sight, they seem to involve a similar [NP PP] structure. This contrast is not surprising, because it also appears in main clauses. Whatever accounts for (i) versus (ii) also accounts for (7) versus (11a):

- (i) John<sub>i</sub> saw some relatives of his<sub>i</sub>  
 (ii) \* Mary<sub>i</sub> told a story about her<sub>i</sub>



- b. *El [ interés del uno en el otro<sub>i</sub>][que mostraron María y Juan<sub>i</sub> e]*  
 the interest of one in the other that showed María and Juan
- c. *Vi a [ los dos pacientes][que visitó cada médico e hoy]*  
 saw to the two patients that visited each doctor today  
 (cada medico >> dos pacientes)
- d. \**La [historia sobre María<sub>i</sub>][que contó ella<sub>i</sub> e]*  
 the story about María that told she
- e. ??*La [historia sobre María<sub>i</sub>][que piensa ella<sub>i</sub> [que Pedro contó e]]<sup>12</sup>*  
 the story about María that thinks she that Pedro told

The generalization seems to be that Spanish SV relatives can be accounted for in the same way as their English counterparts, in terms of a mixed matching/raising analysis, whereas VS relatives call for a pure raising analysis. The two crucial examples are (13a) and (13e), where the gap contains not an impoverished copy of the head, but a full one. This is evidenced by the binding theory violations they induce, compared to (12a) and (12e). The question, again, is how SV inversion can be related to these paradigms.

#### 2.4A revision of the double-headed analysis

The conclusion of the last section—that SV relatives can receive a mixed matching/raising analysis while VS relatives are only derivable through a raising analysis—is descriptively adequate. However, it begs the question of why such a split should exist, and precisely in this way, out of all logical possibilities. In this section, I will introduce some amendments to Sauerland's (1998, 2000) double-headed analysis that will allow us to establish a direct correlation between the matching analysis and SV order on the one hand and the raising analysis and VS order on the other.

First, I retain the assumption that every relative clause contains two head positions. This is an intuitively natural way to think about relatives. A relative clause is essentially a construction in which a given NP is playing a role in two

<sup>12</sup> Admittedly, as one referee pointed out, the contrast between (12e) and (13e) is fairly subtle. To my ear, it is easier to perceive if the bare pronoun is replaced by the emphatic form *ella misma* 'she herself,' which forces reference to *María* rather than to some other woman in the discourse.

(i) *La historia sobre María que ella misma piensa que Pedro contó*  
 the story about María that she herself thinks that Pedro told

(ii) ??/\*? *La historia sobre María que piensa ella misma que Pedro contó*

To complete this point, note in the following dialogue that it is independently possible to have *ella misma* in postverbal position.

(iii) ✓ ¿Y Pedro? – *María dijo que ella misma lo había asesinado*  
 and Pedro? María said that she herself him had murdered

(iv) ✓ ¿Y Pedro? – *María dijo que lo había asesinado ella misma*

clauses at the same time. There are two ways to formalize this observation. One is implicit in Kayne's (1994) raising analysis: There is, effectively, a single NP, and it somehow comes to play a role in both clauses. The other possibility is implicit in both Chomsky's (1977) empty operator and Platero's (1974) and Sauerland's (1998, 2000) double-headed analysis: There are two NPs, one per clause, and one of them is not pronounced.

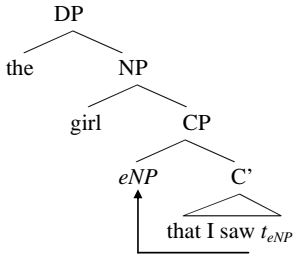
Notice, though, that what I assume is that there are two head positions. At this point, I depart from Sauerland's (1998, 2000) line of analysis and assume that there is only one 'real' head NP, which may occupy either of the two head positions, but, by definition, only one. The other position is occupied by an element I will refer to as '*eNP*.' The purpose of this element is to act as a placeholder of the 'real' head in the head position the latter is not occupying. It is important to stress that I do not place any restrictions on the placement of these two elements. That is, it is possible to have the 'real' head in the clause-external position and *eNP* in the clause-internal one, and vice versa. The reasons behind this assumption will be made clear in a later section.

What is *eNP*? As I said, I assume it is a duplication of the 'real' head NP, with complex internal structure.<sup>13</sup> I will also assume that it has no reference of its own, but rather it has to assume whatever reference the 'real' head has. Nonetheless, it contains a small set of formal features, such as Case,  $\phi$ , Animacy, and so on. In English and Spanish, it is phonetically null, but, in principle, nothing prevents it from being realized in other languages, for instance, as a resumptive pronoun, or even as a duplication of the 'real' head (see fn. 12).

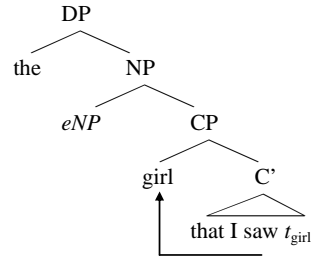
Given these assumptions, we can construct a unique structure for relative clauses that allows for two different ways of building it up. I will continue to refer to these two options as 'matching' (14) and 'raising' (15) relatives, although I want to stress that these structures do not formally correspond to the structures presented in (5) and (6), respectively. I only maintain the terminology for the sake of simplicity.

<sup>13</sup> The idea of *eNP* being syntactically articulate might seem strange at first thought. However, see Grosu's (2003) analysis of free relatives, which he argues are CPs adjoined to a null external head consisting of a number of projections. My analysis goes a little beyond Grosu's in that, while the structure of his external head is invariant, I assume that the structure of *eNP* varies so as to mimic the structure of the 'real' head. Interestingly, there seems to be empirical evidence in favor of this hypothesis. Susagna Muntaña (personal communication, n.d.) informs me that Spanish children sometimes produce sentences of the form *el gato que he visto el gato*, literally, "the cat that I've seen the cat." In other words, the head is duplicated inside the relative. As will become clear in the upcoming discussion, this can be easily accounted for under the present hypothesis. However, I have not yet looked in detail at this kind of data, therefore I will not discuss them further in this chapter.

(14) Matching relatives:



(15) Raising relatives:



Semantically, (14) and (15) are equivalent.<sup>14</sup> Virtually all the analyses of relatives I know of posit movement of some element from the gap position to SpecCP with the purpose of creating an operator-variable chain.<sup>15</sup> In (14), movement of *eNP* fulfills the same role as operator movement in the traditional matching analysis. The variable is the trace/copy of *eNP*, and the binder is *eNP* itself. The same holds for (15), with the difference that it is *girl* that creates and binds the variable. Therefore, the intersection between the relative and the head takes place entirely inside CP.

### 3. *Inversion in Spanish*

#### 3.1 *Obligatory inversion in Spanish*

As we have seen, inversion is linked to reconstruction in Spanish. I take this to suggest that SV inversion in relatives is not a free, 'stylistic' process. Rather, it seems as though it interacts with some other operations. To extend this line of thought, let us look at the different contexts in which inversion is obligatory in Spanish, as shown in (16) through (18).<sup>16</sup>

<sup>14</sup> Thanks to Crit Cremers and Alex Grosu for clarifications and discussion on this point.

<sup>15</sup> A notable exception is Adger and Ramchand (2002), who argue that, in Gaelic languages, lambda abstraction is established between the relative complementizer and a pronominal element in the gap position via the Agree operation.

<sup>16</sup> In this chapter, I only try to derive some generalizations about inversion. I am not committed to how inversion should actually be analyzed. In particular, I believe that the proposal I present here is compatible with a T-to-C movement analysis of inversion as well as with a subject-in-VP approach.

- (16) Matrix *wh* questions:<sup>17</sup>
- a. *¿Qué ha visto Juan?*  
what has seen Juan
  - b. \**¿Qué Juan ha visto?*
- (17) Embedded *wh* questions:
- a. *Me pregunto [qué ha visto Juan]*  
I wonder what has seen Juan
  - b. \**Me pregunto [qué Juan ha visto]*
- (18) Focus fronting:
- a. Un Mundo Feliz *ha leído Juan* (y no Rebelión en la Granja)  
*Brave New World* has read Juan and not *Animal Farm*
  - b. ?\* Un Mundo Feliz *Juan ha leído* (y no Rebelión en la Granja)

On the other hand, inversion is not obligatory in the contexts in (19)-(22):

- (19) Declarative sentences:
- a. (*Esta mañana*) *ha leído Juan el periódico.*  
this morning has read Juan the newspaper
  - b. (*Esta mañana*) *Juan ha leído el periódico.*
- (20) In situ questions:
- a. ??(*María piensa*) *que ha leído Juan qué*  
María thinks that has read Juan what
  - b. (*María piensa*) *que Juan ha leído qué*
- (21) Topic fronting:<sup>18</sup>
- a. Un Mundo Feliz # *lo ha leído Juan.*  
*Brave New World* it has read Juan
  - b. Un Mundo Feliz # *Juan lo ha leído.*
- (22) Successive cyclic *wh* movement:<sup>19</sup>
- a. *¿Qué piensa María [ que ha leído Juan t]?*  
what thinks María that has read Juan
  - b. *¿Qué piensa María [que Juan ha leído t]?*

How can we set both groups apart? Following Rizzi (1997, 2001), I assume that (16)-(18) have quantificational force, that is, they contain an operator-

<sup>17</sup> I am aware of the fact that the obligatoriness of inversion with different types of *wh* words shows a great deal of dialectal variation (Bakovic 1998; cf. also Suñer 1994). The data reported here come from my dialect of northern Spain, where inversion seems to be obligatory with all *wh* words.

<sup>18</sup> Following Rizzi (1997), I will assume that foci differ from topics in that the former show an intonational break and a clitic doubling the topic.

<sup>19</sup> Torrego (1984) marks sentences like (22b) as ungrammatical. However, my informants (and myself) find these examples perfectly acceptable, though it might be true that, in a minimal pair, (22a) might be preferred to (22b). In any event, I believe Torrego's judgments could be accommodated with a minimal revision of (24).

variable chain, whereas this is not the case in (19)-(22). Let us formulate this as (23):

- (23) The inversion generalization (preliminary version):  
 Inversion is obligatory if the clause in question contains an operator-variable chain.  
 Otherwise, it is not.

This statement, as such, runs into problems. To begin with, it would be possible to argue that in situ questions (20) involve covert movement of the *wh* word. Similarly, following Cinque (1990), one might object that in CLLD constructions like (21) there is empty operator movement, the topic being base generated in its surface position. Finally, we might also wonder why inversion is not obligatory in the intermediate landing sites of *wh* movement in (22), given that these positions are links of an operator-variable chain. For these reasons, let us reformulate (23) as (24):

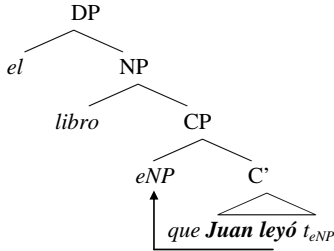
- (24) The inversion generalization (definitive version):  
 Inversion is obligatory if  
 a. the clause in question contains an operator-variable chain  
 AND  
 b. the head of the chain has some phonological content.

At present, I cannot derive this generalization from anything. However, I will show that (24) is useful to account for the relative clause data presented earlier. Therefore, for the time being, let us accept it as I have just phrased it.

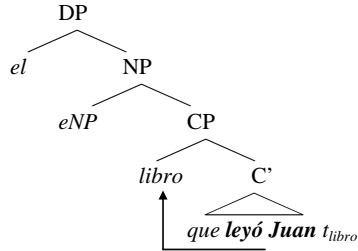
### 3.2 *Inversion in relatives*

Relative clauses also contain an operator-variable chain. Therefore, we can assume that they also abide by (24). What (24) says is, roughly, that overt A-bar movement causes inversion, as opposed to covert—or lack of—A-bar movement. Recall that, according to the hypothesis presented in section 2, both options are available in relative clauses. If the ‘real’ head is inside the relative CP, its raising to SpecCP would be an instance of overt A-bar movement, and would trigger inversion. On the other hand, movement of *eNP* will not trigger inversion, since it has no phonological matrix. Thus, (14) and (15) have the structures (25) and (26) in Spanish:

(25) Matching relatives:



(26) Raising relatives:



Notice that this hypothesis requires having an empty double of the ‘real’ head (namely, *eNP*) from the very beginning, rather than as a consequence of deletion, as Sauerland (1998, 2000) proposes. The reason is that, by definition, PF deletion applies to a finished derivation. As a consequence, if the two heads were phonetically realized, inversion would always be triggered. This would happen irrespectively of which of the heads is deleted later on. That is, under this system, the mere existence of SV relatives favors having an unpronounced second head that is generated as such, rather than derived through deletion at a later stage.

#### 4. *Back to the asymmetries*

##### 4.1 (Anti)reconstruction of the relative clause head

The structures in (25) and (26) amount to saying that SV relatives have an impoverished copy of the ‘real’ head in the gap position, whereas VS relatives have a full one. This derives the asymmetries in (12) and (13). Let us begin with the former. Examples (12a), (12d), and (12e) fall out from assuming that the gap position contains a pronoun instead of the R-expression present in the ‘real’ head. The only requirement, then, is that the pronoun and its binder are not in the same binding domain. This condition is met by (12a) and (12e), but not by (12d), hence the grammaticality of the former and the ungrammaticality of the latter.

Examples (12b) and (12c) require further discussion. In (12b), the anaphor in the ‘real’ head is properly bound, even though the system proposed here implies that the gap only contains an impoverished copy of the ‘real’ head.<sup>20</sup> Thus, it seems as though we have a paradox here. The solution to this problem relies on how much of an ‘impoverished copy’ *eNP* is. Recall that I assumed earlier (see fn. 12) that *eNP* mimics the syntactic structure of the ‘real’ head. By ‘impoverished,’ I mean that *eNP* contains no referential elements, as was

<sup>20</sup> Thanks to Jairo Nunes (personal communication, n.d.) for raising this point.

proposed earlier. That is, the R-expressions of the ‘real’ head correspond to pronouns in *eNP*, which is the same effect as VC (though see the next section for an argument of why this hypothesis is preferable to VC). One way to derive this state of affairs is to assume that Full Interpretation bans the presence of unpronounced R-expressions, since the information they carry would not be recoverable. Therefore, the only elements admitted in such contexts would be pronouns and anaphors, which take their reference from an appropriate antecedent. The consequence is that an anaphor in the ‘real’ head would also correspond to an anaphor in *eNP*. In this way, we derive the grammaticality of (12b). A similar treatment can be applied to (12c). Since the gap contains an element with the same syntactic structure as the ‘real’ head, wide scope of *cada médico* ‘every doctor’ follows in the same way as anaphor binding in (12b).

On the other hand, the VS relatives in (13) have been argued to contain a full copy of the ‘real’ head in the gap position. In other words, whatever is in the ‘real’ head will also be present in the gap. If the ‘real’ head contains an R-expression, as in (13a), (13d), and (13e), we predict that the gap cannot be c-commanded by any element that qualifies as a binder for the R-expression, lest we have a Condition C violation. As we can see, the prediction is fulfilled. Anaphor binding (13b) and wide scope of *cada médico* ‘every doctor’ (13c) also follow without any further stipulation, since all necessary elements (the anaphor in the former case and *dos pacientes* ‘two patients’ in the latter) are present in the gap position as a result of the movement operation.

#### 4.2 (Anti)reconstruction of the entire relative clause

The observation this chapter started off with is that, in Spanish, a relative contained in a fronted *wh* phrase must reconstruct to the base position if it displays VS order, but not otherwise. In this section, I try to derive this generalization from the analysis I developed earlier. Given that I have hypothesized that VS order obtains whenever the ‘real’ head occupies the clause-internal position, what we need is a mechanism that ensures that the relative CP will adjoin cyclically in this situation, but not if the ‘real’ head is in the clause-external position.

Notice, to begin with, that the relation between the ‘real’ head and *eNP* is asymmetric, in the sense that the presence of *eNP* entails the presence of another NP (the ‘real’ head) that it can be related to, whereas a regular NP does not entail the presence of *eNP*. The latter case would simply result in an NP without any relative CP adjoined to it. Starting from this assumption, let us propose that insertion of *eNP* must be followed as soon as possible by insertion of the ‘real’ head, but not the other way around. Now consider a case in which

*eNP* is in the clause-external position. Obviously enough, the ‘real’ head will be in the clause-internal position. This is precisely the situation where cyclic insertion of the relative CP is forced by the assumptions I have just laid out. Once *eNP* is introduced in the derivation, the ‘real’ head must also be introduced. But, since the ‘real’ head is embedded inside the relative CP, the only way to introduce the former is to adjoin the entire CP to *eNP*. In order to satisfy the ‘as-soon-as-possible’ requirement, adjunction must take place before any other operations. As a consequence, we derive the observation that VS relatives are inserted cyclically.

Consider, on the other hand, a situation where it is the ‘real’ head that occupies the clause-external position. Since the presence of the NP that constitutes the ‘real’ head does not entail the presence of *eNP*, there is nothing forcing cyclic adjunction of the relative CP. Instead, adjunction can be delayed until a later point, when the head DP has undergone movement operations.<sup>21</sup>

Notice finally that this analysis relies on having something like *eNP* generated as such, rather than as a result of deletion, as Sauerland (1998, 2000) proposes. If we adopted Sauerland’s hypothesis, there would be no obvious way to force cyclic adjunction in some cases but not in others. On the other hand, if we assume *eNP* is an element of the numeration in its own right, the asymmetry can be easily derived. It follows from this plausible assumption that the relation between the ‘real’ head and *eNP* is asymmetric.

## 5. Open questions

### 5.1 English

In English, inversion is banned from embedded contexts. Therefore, we cannot use the same test as in Spanish to determine which one of the two structures, (14) or (15), is being used, or even if either structure is used at all. On the one hand, it seems as if restrictive relatives could be fully covered if only (14) was used in English. On the other hand, the null hypothesis is that English, like Spanish, can resort to both structures. At present, I do not know of any conclusive evidence in favor of either option (though see Sauerland 1998, 2000; Aoun & Li 2003, for discussion).

---

<sup>21</sup> A related question is, when *eNP* is in the clause-internal position, why is cyclic adjunction not forced in order to relate it to the clause-external ‘real’ head? The answer capitalizes on the assumption that insertion of the ‘real’ head is required only when *eNP* enters the derivation. Therefore, we can propose that, in this situation, the relative CP containing *eNP* is not constructed in parallel. Rather, it would not be derived until it is time for it to be adjoined to the ‘real’ head.



### 5.2 *Relative and resumptive pronouns*

In Spanish, *that* relatives are used for subject and direct object relativization, that is, the two grammatical functions that do not require a preposition. In other cases, relative pronouns are used. Although the judgments are slippery, it seems that inversion has the same effects in sentences with relative pronouns, as in (27):

- (27) a. *El [ amigo de María<sub>i</sub>] [con el que ella<sub>i</sub> ha discutido t]*  
           the friend of María with the that she has argued  
       b. ?? *El [amigo de María<sub>i</sub>] [con el que ha discutido ella<sub>i</sub> t]*

However, it is not obvious at all how to construct these sentences under the perspective of the theory presented here. The toughest problem comes from the variety of forms available for sentences with relative pronouns (some properties of pronouns are discussed by Brucart 1992), shown in (28):

- (28) a. P + (D) + *que<sub>C</sub>* → *la chica con (la) que Juan ha hablado*  
           the girl with the that Juan has talked  
       b. P + *quien<sub>D</sub>* → *la chica con quien Juan ha hablado*  
           the girl with which Juan has spoken  
       c. P + D + *cual* → *la chica con la cual Juan ha hablado*  
           the girl with the *cual* Juan has spoken

There is also a fourth type, which involves no relative pronoun, but resumption of the Hebrew type, that is, in non-island contexts (cf., e.g., Aoun, Choueiri, & Hornstein 2001 for real vs. apparent resumption in Lebanese Arabic), as in (29):

- (29) *La persona que los apuntes son suyos puede pasar a recogerlos*  
       the person that the class-notes are his/hers can come to pick-up-CL  
       ‘‘The person who owns the class notes can come to pick them up.’’<sup>22</sup>

One way to analyze these resumptive pronouns would be to say that they are simply spelled-out copies of *eNP*. In the best of worlds, the difference between languages with and without resumption could be reducible to a parametric difference between spelling *eNP* out or not. One thing that does not follow from the hypothesis presented here is the fact that resumptive pronouns always take place in clause-internal position. That is, given the possibility for

<sup>22</sup> This sentence was heard in a radio program; the speaker had found a folder with class notes and was looking for the owner. Sentences like this are almost exclusively produced in spoken contexts. In written communication, the predominant form is the relative pronoun *cuyo* ‘‘whose.’’

*eNP* to appear in the clause-external head position, what prevents structures like (30b)?

- (30) a. [DP...NP [CP ...resumptive pronoun...]]  
 b. \* [DP ...resumptive pronoun [CP ...NP...]]

### 5.3 Correlatives and internally headed relatives

Actually, the pattern in (30b) recalls a correlative clause, as in Hindi. Unfortunately, typological studies (de Vries 2002) reveal that, in correlative clauses, the relative CP nearly always precedes the pronoun, which does not fall out from the theory presented here at all. However, the theory, in principle, provides the elements to derive correlative clauses in a similar way to ‘regular’ relative clauses. In any event, I must leave this issue open.

On the other hand, internally headed relative clauses are explained fairly easily. Following the idea of Platero (1974) schematized in (9) (and taken up again recently by Kayne 1994 and Bianchi 1999, 2000), such a clause would simply be one in which the copy of the ‘real’ head in a structure like (14) or (18) that is spelled out is the one in the gap position, and not in SpecCP.

To sum up, one of the strongest points of the theory presented in this chapter is that it provides a means to derive all kinds of relatives—both internally and externally headed, with and without resumption, and correlatives—in a strikingly similar way. There are still many problems remaining at this stage, which will hopefully be solved in the future.

## REFERENCES

- Adger, David & Gillian Ramchand. 2002. “Moving Relatives Can Be Avoided (If We Choose to AGREE)”. Paper presented at the On Wh- Movement workshop, Leiden-Utrecht, December 2002.
- Alexiadou, Artemis, Paul Law, André Meinunger, & Chris Wilder. 2000. *The Syntax of Relative Clauses*. Amsterdam & Philadelphia: John Benjamins.
- Aoun, Joseph, Lina Choueiri, & Norbert Hornstein. 2001. “Resumption, Movement, and Derivational Economy”. *Linguistic Inquiry* 32.371-403.
- Aoun, Joseph & Yen-hui Li. 2003. *Essays on the Derivational and Representational Nature of the Grammar: The diversity of wh- constructions*. Cambridge, MA: MIT Press.
- Bakovic, Eric. 1998. “Inversion and Optimality in Spanish”. *Is the Best Good Enough?* ed. by Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis, & David Pesetsky, 35-58. Cambridge, MA: MIT Press.
- Bhatt, Rajesh. 2002. “The Raising Analysis of Relative Clauses: Evidence from adjectival modification”. *Natural Language Semantics* 10.43-90.
- Bianchi, Valentina. 1999. *Consequences of Antisymmetry: Headed relative clauses*. Berlin: de Gruyter.

- , 2000. "The Raising Analysis of Relative Clauses: A reply to Borsley". *Linguistic Inquiry* 31.123-140.
- Bobaljik, Jonathan. 2002. "A-Chains at the PF Interface: Copies and 'covert' movement". *Natural Language and Linguistic Theory* 20.197-267.
- Bresnan, Joan. 1973. "Syntax of the Comparative Clause Construction in English". *Linguistic Inquiry* 4.275-343.
- Brucart, José. 1992. "Some Asymmetries in the Functioning of Relative Pronouns in Spanish". *Catalan Working Papers in Linguistics* 7.113-143.
- Chomsky, Noam. 1977. "On Wh- Movement". *Formal Syntax* ed. by Peter Culicover, Thomas Wasow, & Adrian Akmajian, 71-132. New York: Academic Press.
- , 1993. "A Minimalist Program for Linguistic Theory". *The View from Building 20* ed. by Kenneth Hale & Samuel J. Keyser, 1-52. Cambridge, MA: MIT Press.
- Cinque, Guglielmo. 1990. *Types of A' Dependencies*. Cambridge, MA: MIT Press.
- de Vries, Mark. 2002. *The Syntax of Relativization*. Ph.D. dissertation, University of Amsterdam.
- Epstein, Samuel, Erich Groat, Ruriko Kawashima, & Hisatsugu Kitahara. 1998. *A Derivational Approach to Syntactic Relations*. Oxford: Oxford University Press.
- Fiengo, Robert & Robert May. 1994. *Indices and Identity*. Cambridge, MA: MIT Press.
- Fox, Danny. 1999. "Reconstruction, Binding Theory, and the Interpretation of Chains". *Linguistic Inquiry* 30.157-196.
- Grosu, Alexander. 2003. "A Unified Theory of Standard and Transparent Free Relatives". *Natural Language and Linguistic Theory* 21.247-331.
- Gutiérrez Bravo, Rodrigo. 2002. *Structural Markedness and Syntactic Structure: A study of word order and the left periphery in Mexican Spanish*. Ph.D. dissertation, University of California, Santa Cruz.
- , 2003. "Subject Inversion in Spanish Relative Clauses". Paper presented at Going Romance 17, University of Nijmegen, November 2003.
- Heycock, Caroline. 1995. "Asymmetries in Reconstruction". *Linguistic Inquiry* 26.547-570.
- Kayne, Richard. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Lebeaux, David. 1988. *Language Acquisition and the Form of the Grammar*. Ph.D. dissertation, University of Massachusetts, Amherst.
- , 1992. "Relative Clauses, Licensing, and the Nature of the Derivation". *Syntax and Semantics* 25.209-239.
- Platero, Paul. 1974. "The Navajo Relative Clause". *International Journal of American Linguistics* 40.202-246.
- Platzack, Christer. 2000. "A Complement-of-N Analysis of Restrictive and Non-restrictive Relatives: The case of Swedish". *The Syntax of Relative Clauses* ed. by Artemis Alexiadou, Paul Law, Andre Meinunger, & Chris Wilder, 265-308. Amsterdam & Philadelphia: John Benjamins.
- Rizzi, Luigi. 1997. "The Fine Structure of the Left Periphery". *Elements of Grammar* ed. by Liliane Haegeman, 281-335. Dordrecht: Kluwer.
- , 2001. "Locality and Left Periphery". Ms., University of Siena.
- Sauerland, Uli. 1998. *The Meaning of Chains*. Ph.D. dissertation, Massachusetts Institute of Technology.
- , 2000. "Two Structures for English Restrictive Relative Clauses". Ms., University of Tübingen.
- Stepanov, Arthur. 2001a. *Cyclic Domains in Syntactic Theory*. Ph.D. dissertation, University of Connecticut.

- , 2001b. "Late Insertion and Minimalist Phrase Structure". *Syntax* 4.94-125.
- Suñer, Margarita. 1994. "Verb Movement and the Licensing of Argumental Wh- Phrases in Spanish". *Natural Language and Linguistic Theory* 12.335-372.
- Torrego, Esther. 1984. "On Inversion in Spanish and Some of Its Effects". *Linguistic Inquiry* 15.103-129.
- vanden Wyngaerd, Guido & Jan-Wouter Zwart. 1991. "Reconstruction and Vehicle Change". *Linguistics in the Netherlands* 1991.151-160.
- Vergnaud, Jean-Roger. 1974. *French Relative Clauses*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Williams, Edwin & Henk van Riemsdijk. 1981. "NP Structure". *The Linguistic Review* 1.171-218.



# INFINITIVAL COMPLEMENT CONSTRUCTIONS IN SPANISH A CONSTRUCTION GRAMMAR APPROACH\*

JIYOUNG YOON  
*University of North Texas*

## 0. *Introduction*

Previous approaches to infinitival complement constructions have been mainly syntactic (Bošković 1997; Chomsky & Lasnik 1993; Stowell 1982, among others), with some semantic-oriented studies (Clements 1992; Givón 1990; Wierzbicka 1988, among others).<sup>1</sup> Most syntactic studies have tried to determine the categorial status of the infinitival complements to explain different syntactic behaviors found in different types of infinitival complement constructions.<sup>2</sup> To carry out such an analysis, they have relied on the idea that the main verb is primarily responsible for determining the overall argument structure of the sentence. This view is well expressed in Chomsky's Projection Principle (1981:29), according to which lexical items project their syntactic properties into a sentence. For instance, *pensar* "think" would be assumed to 'select' three syntactic subcategories: a sentential complement as in (1), an infinitival complement as in (2), or a prepositional phrase (preposition accompanied by an infinitive) as in (3):

- (1) *La gente piensa que fui quien decidió.* (RAE)<sup>3</sup>  
"The people think/believe that I was the one who decided."

---

\* I would like to thank two anonymous reviewers for their detailed and valuable comments. I am also grateful to Adele Goldberg for helping me clarify some theoretical aspects of Construction Grammar. All errors are solely my own.

<sup>1</sup> Some semantic-oriented studies sought to find the meaning of infinitival complements compared to other complement types such as gerunds and sentential complements in languages like English, whereas very few semantic studies have focused on lexical aspects of infinitives or the diachronic evolution of infinitival complement constructions.

<sup>2</sup> See Bošković (1997) for more discussion about control verbs such as *want*-class verbs and *try*-class verbs taking IP or CP infinitival complements.

<sup>3</sup> Most data in this study were taken from the Real Academia Española database at <http://www.rae.es> (indicated as RAE) and Spanish newspapers *Hispanidad* (HP, March 21, 2003) and *El país* (EP, August 26, 1999). All other examples were created by the author, and were judged by three native Spanish speakers for grammaticality/acceptability.

- (2) *El grupo público piensa estar presente en todas aquellas comunidades...* (RAE)  
 “The public group intends to be present in all those communities...”
- (3) *Hoy la gente sólo piensa en llegar a casa con su coche.* (RAE)  
 “These days people only think of getting home in their car.”

In an infinitival complement construction as in (2), *pensar* “think” is construed as “to intend or plan to (do something)” (a volition verb) while the same verb with the sentential complement in (1) is no longer interpreted as an intentional verb, but rather as a cognitive verb such as *creer* “believe.” With the construction [*en* + infinitive] as in (3), it also has a pure cognitive meaning. From a traditional syntactic point of view, the different argument structures with different meanings of the verb shown in (1)-(3) can be accounted for by positing different semantic senses of the verb *pensar* “think,” that is, the meaning of the main verb is responsible for determining its complementation structures. The implication of this proposal is that children as well as learners of a second language acquire and master argument structures for each individual verb on an item-by-item basis.

However, it might be difficult to assume that children actually learn an infinite number of different verbs (or even newly created verbs) and their argument structures based on the input they receive each time they hear a verb occurring in different syntactic structures. That is, it might not be plausible to think that children learn a language without generalizing and categorizing the common semantic and formal relationships found between verbs and their associated semantic and syntactic structures. Instead of positing different senses for the same verb, it is more reasonable to postulate that the language acquisition process occurs by associating the semantics of the verb directly with formal patterns, that is, abstract constructions (Goldberg 1999). For example, the intentional meaning of *pensar* “think” in (2) is obtained not necessarily because the verb itself has this meaning, but because both the verb and the construction in which the verb occurs provide meaning and because the verb receives a different semantic construal in this particular construction, namely, the infinitival complement construction. This idea, proposed in Construction Grammar (Goldberg 1995, 1998, 1999; Lakoff 1986, 1987), claims that each abstract linguistic pattern or construction, which is a pairing of form and meaning, has a meaning of its own.

Postulating meanings of constructions allows us to better account for the examples in (4):

- (4) a. *\*El joven piensa ser inteligente.*  
 “The young man intends to be intelligent.”

- b. *El joven piensa ser ingeniero en el futuro.* (RAE)  
 “The young man intends to be an engineer in the future.”

Example (4a) is not a possible sentence since the infinitival complement *ser inteligente* “be intelligent” refers to inherent properties of the subject that cannot be planned to be achieved. Thus, semantically this example does not fit the meaning of the infinitival complement construction, which I assume to be DESIRE-BECOME, to be discussed in detail in the following sections. The example in (4b), on the other hand, fits the meaning of the infinitival complement construction; thus, it is grammatical. The contrast shown in (4a)-(4b) clearly supports the existence of the infinitival complement construction in Spanish that has its own meaning. I argue that the Construction Grammar approach (Goldberg 1995, 1998, 1999; Lakoff 1986, 1987) better explains Spanish infinitival complement constructions by accounting for semantic as well as syntactic aspects that many formal approaches have attributed to the individual meanings of the verbs.

### 1. *Semantic motivations*

In this section, I provide semantic motivations for positing two infinitival complement constructions in Spanish: the DESIRE-BECOME construction and the ASSESS-STATE construction.

#### 1.1 *Semantics of main verbs*

Spanish verbs appearing in infinitival complement constructions can be clearly divided into two general groups based on the notion of volition (cf. Rudanko 1998): [+volition] such as verbs of desideration, intention, and attempt, and [-volition] such as verbs of emotion, cognition, and declaration, as shown in (5) and (6), respectively.

- (5) [+volition] verbs:
- a. Desideration (*esperar* “hope,” *querer* “want,” *anhelar* “hope,” *preferir* “prefer”):  
*Usted no quiere bailar hasta muy mayor.* (RAE)  
 “You don’t want to dance until (you are) very old.”
  - b. Intention (*intentar* “intend,” *prometer* “promise,” *decidir* “decide,” *resolver* “resolve/decide,” *pensar* “think, intend,” *rechazar* “refuse,” *aceptar* “agree,” *acordar* “agree,” *proponer* “intend,” *pretender* “intend”):  
*Con la rebaja del 6% el Gobierno pretende ahorrar el año que viene 52.000 millones de pesetas...* (EP)  
 “With the reduction of 6%, the government intends to save 52 billion pesetas next year...”



- c. Attempt/effort<sup>4</sup> (*buscar* “seek,” *evitar* “avoid,” *aprobar* “support”):  
*Somos un grupo autónomo que busca obtener pronto la personalidad jurídica.*  
 (HP)  
 “We are an independent group that is seeking to establish our legal representation as soon as possible.”
- (6) [–volition] verbs:
- a. Emotive/factive (*lamentar* “regret,” *odiar* “hate”):  
*...por eso odia ser como su madre sometida...* (RAE)  
 “...therefore she hates being like her oppressed mother...”
- b. Cognition/mental act or belief (*creer* “believe,”<sup>5</sup> *dudar* “doubt,” *reconocer* “acknowledge,” *negar* “deny”):  
*...un 42% de los consultados reconoce tener miedo.* (EP)  
 “...42% of those surveyed acknowledge being afraid.”
- c. Declaration (*declarar* “declare, claim,” *afirmar* “affirm, declare”):  
*El Banco de México declara tener treinta mil millones de dólares en...* (RAE)  
 “The Bank of Mexico claims to have 30 trillion dollars in...”

The semantic distinction between [+volition] and [–volition] in the main verb motivates the postulation of two different types of infinitival complement constructions, which I call ‘DESIRE-BECOME’ and ‘ASSESS-STATE’ constructions, respectively. The central sense DESIRE-BECOME in infinitival complement constructions involving [+volition] verbs comes from the common meaning of such constructions as “a subject desires/intends to accomplish an action or to achieve a certain state.” On the other hand, ASSESS-STATE, which is the central sense of the infinitival complement constructions involving [–volition] verbs, is a common sense found in verb types shown in (6). This central sense can be construed as “one assesses and evaluates one’s situation or state.” The meanings of the two constructions will be discussed in detail in section 2.

<sup>4</sup> Verbs such as *conseguir* “manage to,” *lograr* “succeed,” and *no conseguir* “fail to” arguably belong to this attempt/effort subgroup. For example, *conseguir* can have the sense of “one attempts to do something, and as such, one achieves it.” The only difference between *conseguir* verb types and verbs of attempt such as *buscar* “seek” is that in *conseguir*, when one manages to do something, it means that the attempted action or state actually occurred or was reached (for more discussion about English verbs *manage* and *attempt*, see Wierzbicka 1988). For this reason, I consider that *conseguir* and *lograr* (manage verb type) arguably belong to the subgroup of verbs of attempt, but with an additional semantic trait.

<sup>5</sup> The verb *creer* can take infinitival complements in some dialects, but in other dialects the same verb can take only a sentential complement.

## 1.2 *Semantics of infinitival complements*

1.2.1 *Truth value.* Another difference between the two constructions posited is found in infinitival complements themselves. These two constructions differ in the truth value of the infinitival complement.

Most [+volition] verbs, except for a few verbs of attempt such as *conseguir* “manage to,” take a proposition as their complement, in the sense that the statement denoted by the complement cannot have a truth value in a given situation (Castañeda 1975; Clements 1992). For example, in a sentence like *Pretendo hablar con el profesor esta tarde* “I intend to speak to the professor this afternoon,” truth value is irrelevant because the action of speaking to the professor is a future intention. One might speak of the truth value in terms of whether or not I actually intend to speak with the professor, but not in terms of whether or not I have done so.

Verbs of the [-volition] type, on the other hand, involve a proposition because the statement denoted by the complement can have a truth value: A predicate-containing utterance should be either true or false in a given situation (Castañeda 1975; Clements 1992). For instance, in a sentence like *dudo estar embarazada* “I doubt being pregnant,” if it is true that I doubt, then it should be either true or false that I am pregnant.

1.2.2 *Lexical aspect.* When further examining the semantics of the infinitival complements with [+volition] verbs in (5) and [-volition] verbs in (6), we notice that the lexical aspect of the infinitival complement verbs are not uniform between the two constructions involving both verbs. In the DESIRE-BECOME construction involving [+volition] main verbs, the lexical aspects denoted by infinitival complements are mostly achievement (e.g., *obtener* “obtain”), accomplishment (e.g., *ahorrar* “save (money)”), and activity (e.g., *bailar* “dance”). Stative infinitival complements, in contrast, are allowed with some semantic restrictions. Infinitival complements involved in all [+volition] verbs denote states that one can control. Examples are given in (7).

- (7) a. *...quiere poseer tu amor en exclusiva...* (RAE)  
 “...he wants to possess your love exclusively...”  
 b. *Alonso intenta ser optimista dentro del difícil momento...* (RAE)  
 “Alonso intends to be optimistic in the difficult moment...”  
 c. *Se busca estar delgada, y se controla el apetito.* (RAE)  
 “One seeks to be slim, and one controls the appetite.”

However, if the state denoted by infinitival complements cannot be controlled, such as being intelligent or being a woman, only verbs of desire in which the volition of the subject is relatively low are likely to be allowed, as shown in

(8a). In contrast, the same stative infinitival complements yield pragmatically unacceptable sentences if the volition of the main verb is stronger than desire—if it is a verb of attempt, for example, as in (8b):<sup>6</sup>

- (8) a. *Yo quiero ser mujer.* (RAE)  
 “I want to be a woman.”  
 b. *??Intento/le prometi/evito ser mujer.*  
 “(lit.) I intend/promised him/avoid to be/being a woman.”

This inconsistent behavior stems partially from the fact that *querer* in (8a) is a verb of desideration in which a speaker can merely express his/her desire to be in a certain state without necessarily implying that he/she made efforts to reach such a state. Thus, one can want or desire to be a woman, for instance, but it would implausible and pragmatically odd for one to actually intend or seek to be a woman, since one’s sex is not something that can be controlled.<sup>7</sup> Here, we can say that sentence (8b) does not fit the meaning of the construction DESIRE-BECOME, and consequently it is unacceptable.

On the other hand, when analyzing the lexical aspects of the infinitival complements involved in [–volition] verbs in (6), we find that a relatively narrow range of lexical aspects is allowed. In general, the infinitival complement involved in [–volition] verbs is stative (e.g., *ser como su madre* “be like her mother,” *tener miedo* “be afraid,” *tener treinta mil millones de dólares* “have thirty trillion dollars,” etc.). If the lexical aspect of the bare infinitival complement is eventive rather than stative, sentences become pragmatically odd. This further supports my postulation that the meaning for infinitival complement constructions involving [+volition] verbs should contain STATE.

- (9) *??Dudo estudiar mucho.*  
 “(lit.) I doubt to study a lot.”  
 (10) *??Reconozco/declaro {dejar el trabajo/casarme con Mario}.*  
 “(lit.) I acknowledge to quit the job/to marry Mario.”  
 (11) *??Lamento {comprar la computadora cara/bailar tan mal}.*<sup>8</sup>  
 “(lit.) I regret to buy the expensive computer/to dance too badly.”

<sup>6</sup> See (12) for the volition scale.

<sup>7</sup> Such a situation is pragmatically odd, even if in a forced interpretation, the sentence in (8b) is construed as “I intend to be a woman by converting my sex through surgery.” Still, this interpretation is not as natural as in the sentence with the verb *querer* “want.”

<sup>8</sup> The sentence *lamento bailar tan mal* “(lit.) I regret to dance too badly” can be arguably interpreted as “(lit.) I regret to be a bad dancer,” that is, as a stative characteristic of a person. In such a case, the sentence sounds better.

In the following section, after defining the two constructions in terms of Construction Grammar, I argue that these two types of constructions share the same form, but with different meanings.

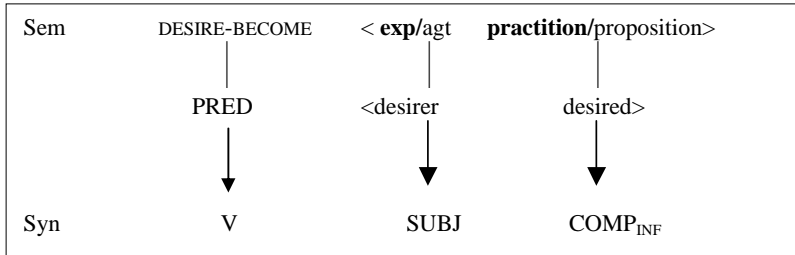
## 2. *The two infinitival complement constructions*

### 2.1 *DESIRE-BECOME construction*

Verbs of [+volition] are typically analyzed as verbs involving so-called ‘subject control’ in the traditional sense, which implies subjects’ desire and inclination to do or not to do something. This verb type can be divided into further subgroups of semantic classes based on the degree of strength of the volition involved in the subject (cf. Givón 1990; Rudanko 1998; Sag & Pollard 1991). They are verbs of desideration expressing subjects’ desire (but with weaker volition than intention and effort verbs) such as *esperar* ‘hope,’ verbs of intention expressing a subject’s intention to do or not to do something such as *prometer* ‘promise,’ and verbs of effort expressing a degree of effort on the part of the agent subject such as *buscar* ‘seek.’ This scale of strength of volition is represented in (12):

(12) desideration < intention < effort

Desideration is weaker than intention in the volition scale since desire is less controllable and realizable than intention, in the sense that one can desire/want to be in a certain state or to do something even if it is something that one cannot control, whereas one generally intends to achieve something that is possible to do and for which one can control the outcome. However, the central meaning residing in all three types of volition verbs is common regardless of the degree of strength of the volition. I have argued that the central sense common to the infinitival complement constructions involving [+volition] verbs is DESIRE-BECOME. I consider desideration as the default meaning in a volition scale (see [12]). I also argue that for other infinitival complement constructions with [+volition] verbs, such as verbs of intention and attempt, the meanings of INTEND-BECOME and ATTEMPT-BECOME, respectively, are inherited from the central sense of DESIRE-BECOME. I return to this relationship between constructions in section 3.

Fig. 1: *DESIRE-BECOME* construction

I posit the structure in Figure 1 for the infinitival complement construction with [+volition] verbs. The box represents the independent status of this construction. The central sense of this construction is indicated as DESIRE-BECOME. The meaning of DESIRE is considered a common semantic trait of all kinds of volition verbs. For example, intention is a promise to oneself or to someone else to intend to achieve a ‘desired’ action or state. Attempting is one’s effort to achieve a ‘desired’ action or state. On the other hand, the meaning of BECOME implies that, by wanting to achieve the intended action or state, the state of not doing something/not being in a certain state may be changed to another state of doing something or being in a certain state (BECOME). The subject of verbs of volition can be either an experiencer if the degree of volition is relatively weak, as in verbs of desire, or an agent if the volition of the subject is strong, as in verbs of attempt (this idea is indicated as <exp/agt> in the box). The infinitival complements can be either *practition* or *proposition* (<practition/proposition>). The boldface indicates which of them is ‘profiled.’<sup>9</sup> In verbs of desideration, the argument role of the subject is experiencer (exp), while the infinitival complement is *practition*, which represents desideration to do or not to do something. The argument role of agent (agt) will be profiled in other volition verb types such as verbs of attempt, in which a subject has volition to accomplish an action or achieve a state, thus being an agent of the sentence. *Proposition*, on the other hand, will be profiled in some verbs of attempt such as *conseguir* “manage to,” since the infinitival complement has a truth value in a given spatio-temporal situation. The bottom tier represents the abbreviated/simple version of the syntax of the

<sup>9</sup> The term ‘profiling’ is used here as it is in Cognitive Grammar, in which it is a type of prominence from the ‘base.’ A well-known example is presented by Langacker (1988:153), who explains that we know that “*aunt* and *niece* contrast semantically by virtue of profiling different participants within the same conceived kinship relation, which functions as their base.”

construction (Syn).<sup>10</sup> V is the verb. SUBJ (subject) and COMP<sub>INF</sub> (infinitival complement) are grammatical roles. In the middle tier, PRED (predicate) represents any verb to be integrated into the construction. The elements <desirer> and <desired> are participant roles for a verb of desideration. The solid line between the argument roles and the verb's participant roles indicates which roles of the construction are obligatorily fused with roles of the verb.

2.2 ASSESS-STATE construction

Along with the DESIRE-BECOME construction, I posit an independent construction involving [-volition] verbs such as cognition, emotive/factive, and declaration verbs. This construction is semantically motivated, as shown in section 1. I argue that this construction differs from the one involving [+volition] verbs in its semantics, even if the syntactic structures in the two constructions appear to be the same. I assume that each of the subtypes of the infinitival complement construction involving [-volition] verbs inherits its semantic and syntactic properties from the abstract superconstruction whose central sense is ASSESS-STATE, as shown in Figure 2.

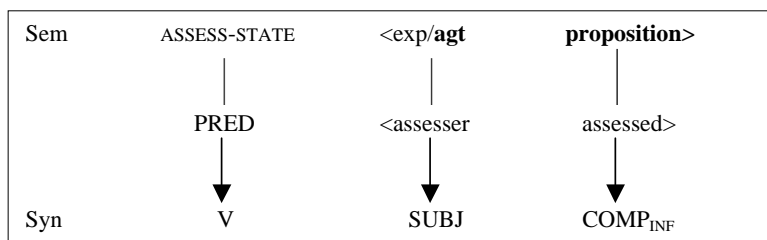


Fig. 2: Abstract ASSESS-STATE construction

Infinitival complement constructions involving emotive/factive, cognition, and declaration verbs are assumed to inherit the meaning and the form of the abstract ASSESS-STATE construction. (I return to this point in the following section.)

The main reasoning for naming the central sense of this construction as ASSESS-STATE is as follows. When one regrets/doubts being poor or when one declares himself/herself to be poor, as illustrated in (13), the common semantics underlying these statements is that one assesses and evaluates one's situation or state. The assessment can be done mentally as in *dudar/creer*

<sup>10</sup> The framework that I follow in this chapter is based on Goldberg's (1995) version of Construction Grammar, not Fillmore and Kay's (1999). Thus, the type of diagram that I use is a relatively simplified version of the more completely specified diagram in Fillmore and Kay.

“doubt/believe” as well as verbally as in *declarar* “declare.” One can also assess a situation by just feeling it, as in *lamentar* “regret.”

- (13) *Lamenta*  
*Duda/creo/reconoce* } *ser pobre.*  
*Declara/afirma*  
 “(lit.) S/he regrets/doubts/believes/acknowledges/claims to be poor.”

In other words, the subject expresses a judgment over a situation or state (rather than a statement of fact), thereby evaluating or assessing it. This point can be confirmed by the unacceptable sentence (14a), in which the subject does not evaluate his/her own situation, but merely conveys a message about it.

- (14) a. ? *La profesora les anuncia ser profesora suplente en la clase de español hoy.*  
 “(lit.) The professor announces to them to be a substitute teacher in the Spanish class today.”  
 b. *La profesora les anuncia que hoy será profesora suplente en la clase de español.*  
 “The professor announces to them that today she will be a substitute teacher in the Spanish class.”

Example (14a) is more acceptable if coded as a finite, sentential complement, as in (14b). The minimal pair in (15) further shows the contrast in the syntactic behaviors between these two verbs.

- (15) a. *No declararon que Juan fuera el asesino.* (subjunctive)  
 “They did not claim that Juan was the murderer.”  
 b. *No anunciaron que iba a haber un congreso en Los Angeles.* (indicative)  
 “They did not announce that a conference was going to be held (lit.: there would be a conference) in Los Angeles.”

In (15a), *no declararon* is interpreted as *no afirmar* “they did not claim,” and thus, the truth value of the infinitival complement, “Juan was the murderer,” cannot be predicted. It can be either true or false that Juan was the murderer; thus the verb in the subordinate clause takes the subjunctive. In contrast, the negation of *anunciar* is not interpreted in the same way as in *no declarar*. In (15b), *no anunciaron* means *no dijeron* “they did not say.” In other words, it is true that a conference was going to be held in Los Angeles, but the subject simply does not say that this is the case. This difference in syntactic behavior is evidence that these two verbs have different syntactic properties, and thus, *anunciar* does not fit the same construction as *declarar*.<sup>11</sup>

<sup>11</sup> This idea is further confirmed by sentences (ia) and (ib), which parallel the contrast in (15):

- (i) a. *No creo que Juan sea mi novio.* (subjunctive)  
 b. *No le digo que Juan es mi novio.* (indicative)

Having argued that the construction under analysis should have ASSESS as its central sense, I now focus on the second element of the central meaning, namely, STATE. Unlike the infinitival complement construction involving [+volition] verbs, the construction involving [-volition] verbs does not code a hypothetical future; rather, two events coded in this construction type should be cotemporal when the infinitives are bare infinitives (i.e., not the form of [*haber* + past participle]). The cotemporality condition is well contrasted in (16), where the adverbial phrase *el próximo año* “next year” is not compatible with the infinitival complement in (16b).

- (16) a. *Una señora madurona que reconoce tener 40 años...* (RAE)  
 “An old lady who acknowledges being 40 years old...”  
 b. \**Una señora madurona que reconoce tener 40 años el próximo año...*  
 “An old lady who acknowledges being 40 years old next year...”

The lexical aspect of the infinitival verb in this construction is stative, as shown earlier in (6). Activities, accomplishments, and achievements are more likely to be disallowed as infinitival complements in this construction, as seen in (9)-(11). One of the ways to make the sentences in (11)-(13) acceptable is to add the auxiliary *haber* “have” to the infinitival verbs to form a perfective aspect, as in *haber llegado* “have arrived,” to denote a past event that has already occurred.

- (17) *Dudo/reconozco/lamento haber dejado el trabajo.*  
 “(lit.) I doubt/acknowledge/regret to have quit the job (I doubt/acknowledge/regret having quit the job).”

The cotemporality condition in [-volition] verbs taking bare infinitival complements is further confirmed by example (18), in which the future-

---

The verb *creer*, as a verb of belief, is a [-volition] verb taking an infinitival complement, like *declarar*. These two verbs take sentential complements introduced by *que* in which the subjunctive mood is employed if sentences are negative, as shown in (15a) and (ia). In contrast, the verb *decir* takes a sentential complement with an indicative verb in negation, as in (ib). These behaviors are exactly parallel to the contrastive behaviors shown in *declarar* versus *anunciar*. The examples in (ii) confirm that *creer* (like *declarar*) can take an infinitival complement, whereas *decir* (exactly like *anunciar*) cannot take an infinitival complement:

- (ii) a. *Creo tener problemas con mi jefe.*  
 b. \**Digo tener problemas con mi jefe.*

All of these parallels demonstrate that *declarar* and *anunciar* cannot occur in the same construction and that only *declarar* can occur in the ASSESS-STATE construction.



oriented state or event is coded as a sentential complement in *dudar* “doubt,” not as an infinitival complement.

- (18) ...*dudo que vaya a apoyarse en la historia de Madrid...* (RAE)  
 “...I doubt that it is going to be supported in the history of Madrid...”

All of these facts support the position that the central sense of the infinitival complement construction involving [–volition] verbs contains the meaning STATE to indicate that the embedded situation does not involve an event but should be a state and that it co-occurs with the main event denoted by the matrix verb. In the following section, I address how Spanish grammar licenses the two constructions through an inheritance network.

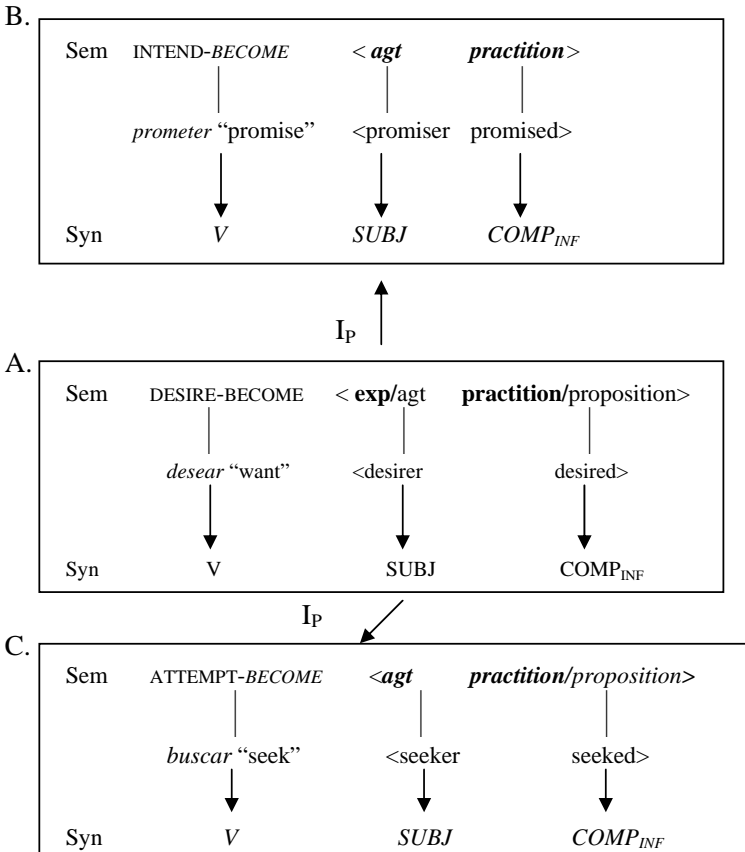


Fig. 3: Polysemy links between [+volition] infinitival complement constructions in Spanish

### 3. *Relations between infinitival complement constructions*

Let us begin by considering the infinitival complement construction involving [+volition] verbs. Recall that [+volition] verbs are divided into different but related semantic groups based on the degree of volition involved. I will take as examples verbs of desideration (e.g., *querer* “want”) and attempt (e.g., *buscar* “seek”) to show the relation between these two constructions.

In verbs of attempt or effort such as *buscar* “seek,” an agent not only has a desire to accomplish a task or to achieve a state, but also acts or makes an effort to accomplish it. Therefore, as illustrated in Figure 3 (A and C), the infinitival complement construction involving verbs of attempt can be regarded as an extension of the dominating infinitival complement construction involving verbs of desideration. All information about syntactic specifications is inherited from the central sense of the dominating construction. The inherited information is indicated in italic letters in the figure.

Following Goldberg (1995), I argue that the infinitival complement construction involving verbs of attempt is licensed by being combined with the infinitival complement construction involving verbs of desideration. These two constructions are combined with an inheritance link, or more specifically, a polysemy link (represented as ‘I<sub>p</sub>’ in Figure 3) like those described by Goldberg. The INTEND-BECOME construction involving verbs of intention as matrix verbs is also argued to inherit the central sense of DESIRE-BECOME by polysemy links, as illustrated in Figure 3 (A and B). Thus the infinitival complement construction involving volition verbs in Spanish can be viewed as a case of ‘constructional polysemy,’ that is to say, that the same form is paired with different but related senses.

Following the same line of argumentation, I propose that inheritance links should also be posited between infinitival complement constructions involving [–volition] verbs such as emotive/factive, cognition, and declaration verbs. The intuition lies in the fact that these constructions are related to each other in their forms and meanings. I argued in the previous section that the central sense that is common to all these constructions is ASSESS-STATE. This class of constructions can be represented as an inheritance network as in Figure 4, in which the central sense ASSESS-STATE is extended to other UTTER ASSESS-STATE, FEEL ASSESS-STATE, and THINK ASSESS-STATE constructions by polysemy links.

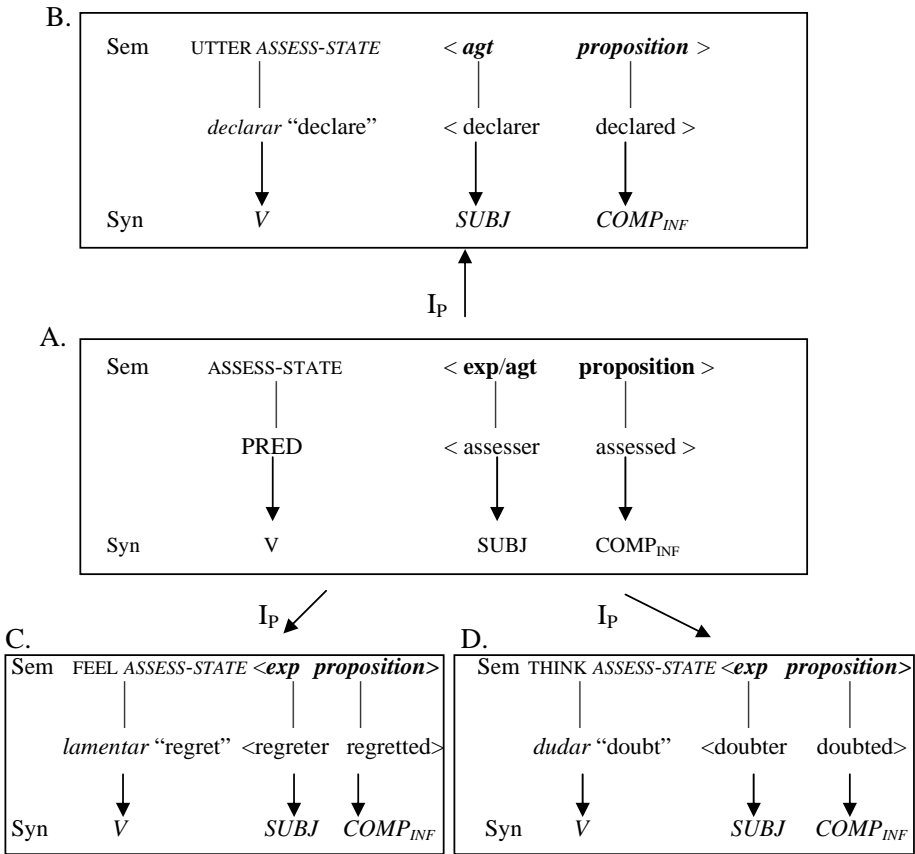


Fig. 4: Polysemy links between [-volition] infinitival complement constructions in Spanish

The analysis of two types of infinitival complement constructions in Spanish, one involving volition of the subject and the other not, brings up the question of what kind of relationship these two constructions have. It is clear that the two constructions have different meanings, but the same structure. I will call a case like this ‘homology of constructions,’ as the form is the same but the meaning is different. Just as ambiguity exists in the case of words, it is reasonable to expect that ambiguity might occur in the case of constructions. The decision to posit two infinitival complement constructions can also be supported by the fact that verbs of volition obligatorily take infinitival complements when the subject of the infinitive refers to the same subject as the matrix verb, as shown in (19). In contrast, [-volition] verbs can take either

infinitival complements or sentential complements (even if many native speakers prefer to use sentential complements with this type of verb) when the subject of the matrix verb refers to the same subject as the infinitive as in (20):

- (19) a. *Quiero ir a tu casa.*  
 “I want to go to your house.”  
 b. \* *Quiero que yo vaya a tu casa.*  
 “(lit.) I want that I go to your house.”
- (20) a. *Mi hermano declaró ser inocente en el juicio.*  
 “(lit.) My brother claimed/declared to be innocent in the court.”  
 b. *Mi hermano declaró que era inocente en el juicio.*  
 “My brother claimed that he was innocent in the court.”

These different syntactic behaviors between the two constructions support the claim that there are two kinds of infinitival complement constructions in Spanish, and here I have argued that they exhibit a homology of construction.

#### 4. *Conclusion*

In the spirit of Construction Grammar, this chapter has argued that the infinitival complement construction in Spanish is associated with a set of systematically related senses, such as desideration, intention, and attempt for the DESIRE-BECOME construction, and declaration, belief, and emotion for the ASSESS-STATE construction. I have claimed that the infinitival complement construction in Spanish can be viewed as a case of ‘constructional polysemy,’ that is to say, that the same form is paired with different but related senses. Spanish grammar licenses these constructions through inheritance networks.

Positing two infinitival complement constructions in Spanish has more explanatory power than merely assuming that the main verb subcategorizes for an infinitival complement. Children acquire complementation structures through a process of generalizing learned instances into patterns. Once constructions are created through the input they receive, this abstract cognitive schema (i.e., construction) can, in turn, facilitate the acquisition process of new verbs and their syntactic structures (Goldberg 1998). The same is true for second language learners. They do not have to memorize all the individual verbs taking infinitival complements, as many Spanish textbooks instruct, without providing further explanations. Rather, by positing the two types of infinitival complement constructions in Spanish, with their own meanings, learners can predict which verb class can fit the meaning and form of the construction to take an infinitival, in the same way that children learn their first language.

It should be noted along these lines that this is why a language may change. Verb classes taking infinitival complements vary by dialect and change over time in the same dialect, as people may innovate and propagate the use of new or existing verbs in a construction in which those verbs did not customarily occur. For example, cognition verbs such as *creer* “believe” do not allow the infinitival complement construction in some dialects, but they do in other dialects. Some native speakers do not accept the infinitival construction with the verb *creer*, as in (21a), but prefer to use a sentential complement as in (21b).

- (21) a.ok/? *Juan cree estar enfermo.*  
 “Juan believes (himself) to be sick.”  
 b. *Juan cree que está enfermo.*  
 “Juan believes that he is sick.”

These are verbs that involve a relatively weaker semantic and syntactic bond between matrix and complement propositions than other types of volition verbs, since they do not tend to form implicative verbs (Givón 1990). It would be interesting to see whether diachronic studies on this construction type would reveal a diachronic extension of meaning and of form, spreading along contiguous portions of a semantic space. Furthermore, a cross-linguistic examination of the infinitival complement construction will help us better understand the coding mechanism of the infinitival complement construction in different languages.

In summary, by recognizing the existence of constructions with different meanings, that is, infinitival complement constructions in Spanish, this chapter suggests that we can avoid the problem of positing different and implausible senses for individual verbs to account for examples like (1)-(3). Rather, this chapter has confirmed the assumption that “linguistic facts that are motivated are neither arbitrary nor predictable. Learners make sense of input forms to the extent that they can identify formal and semantic correspondences among those forms” (Michaelis & Lambrecht 1996:237).

## REFERENCES

- Bošković, Zeljko. 1997. *The Syntax of Nonfinite Complementation: An economy approach*. Cambridge, MA: MIT Press.  
 Castañeda, Hector-Neri. 1975. *Thinking and Doing*. Dordrecht: Reidel.  
 Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

- Chomsky, Noam & Howard Lasnik. 1993. "The Theory of Principles and Parameters". *Syntax: An international handbook of contemporary research* ed. by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld, & Theo Vennemann, vol. I, 506-569. Berlin: de Gruyter.
- Clements, J. Clancy. 1992. "Semantics of Control, Tense Sequencing and Disjoint Reference". *Current Issues in Linguistic Theory 91: Romance languages and modern linguistic theory* ed. by Paul Hirschbühler & Konrad Koerner, 45-56. Amsterdam & Philadelphia: John Benjamins.
- Fillmore, Charles J. & Paul Kay. 1999. "Grammatical Constructions and Linguistic Generalizations: The what's X doing Y? construction". *Language* 75.1-33.
- Givón, T. 1990. *Syntax: A functional-typological introduction*. Amsterdam & Philadelphia: John Benjamins.
- Goldberg, Adele E. 1995. *A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.
- . 1998. "Patterns of Experience in Patterns of Language" *The New Psychology of Language* ed. by Michael Tomasello, 203-219. Mahwah, NJ: Lawrence Erlbaum Associates.
- . 1999. "The Emergence of the Semantics of Argument Structure Constructions", *The Emergence of Language* ed. by Brian MacWhinney, 197-212. Mahwah, NJ: Lawrence Erlbaum Associates.
- Lakoff, George. 1986. "Frame Semantic Control of the Coordinate Structure Constraint". *Papers from the Regional Meetings, Chicago Linguistic Society* 22.152-167.
- . 1987. *Women, Fire, and Dangerous Things: What categories reveal about the mind*. Chicago: University of Chicago Press.
- Langacker, Ronald W. 1988. "Autonomy, Agreement, and Cognitive Grammar". *Papers from the Regional Meetings, Chicago Linguistic Society* 24.147-180.
- Michaelis, Laura A. & Knud Lambrecht. 1996. "Toward a Construction-Based Theory of Language Function: The case of nominal extraposition". *Language* 72.215-247.
- Rudanko, Juhani. 1998. *Change and Continuity in the English Language: Studies on complementation over the past three hundred years*. Lanham, MD: University Press of America.
- Sag, Ivan A. & Carl Pollard. 1991. "An Integrated Theory of Complement Control". *Language* 67.63-113.
- Stowell, Tim. 1982. "The Tense of Infinitives". *Linguistic Inquiry* 13.561-570.
- Wierzbicka, Anna. 1988. *The Semantics of Grammar*. Amsterdam & Philadelphia: John Benjamins.



## INDEX OF TERMS & CONCEPTS

### A

- Accent 323-324
  - Secondary 229
- Adverb
  - Quasi* 344-353
  - Time adverbials 75, 80
- Agent 69, 344-345, 353-354, 356-357
- Agreement 5-11
- Aktionsart 66-67, 69-76, 78-80
- Analogy 270-271
  - Analogical 113
  - see also* Blending
- Aranese
  - see* Gascon
- Argument structure 45-49, 239, 382
- Aspect 66, 69, 72-73, 75-78
  - Aspect phrase 239, 243, 252-254
  - Aspectual head 242-243
  - Aspectual projection(s) 239, 242-243
- Atelicity 66-68, 70-80

### B

- Binding theory 362-370, 373-374
- Blending 261
- Borrowing, Germanic 268

### C

- Case 2, 5-6
  - Case checking 7, 8, 10-11, 239-240
- Catalan 166, 172, 223 (fn. 3), 232, 257, 265, 275-295

- Causatives 244, 246, 248, 253-254
  - Causative alternation 45-47
  - Causer
    - see* Agent
  - Lexical causatives 45-49, 350-351
  - Periphrastic causative *fare* 353-355
- Change of state 45-49, 69, 71, 73, 249-250
- Cleft 205
  - Cleft reduction 206
- Clitic 132, 242, 246
  - Clitic drop 132, 134-136
- Coercion 65, 68, 75-80
- Complementary distribution 300-301
- Complex segments 85, 97-98
- Conjugation 17-19, 23, 34, 36-42, 177-183, 186, 188-189, 191-192, 194, 277-278
  - see also* Paradigm
- Constraints
  - Local conjunction of 309ff.
  - On (dorsal) place of articulation 309
  - On nasal-obstruent sequences 307ff.
  - see also* Faithfulness constraints
- Construction 382, 387-389, 393-394
  - Construction Grammar 48, 382, 395
- Contrast 160, 275, 282-283
- Control relations 384, 387
- CP 3, 5, 7-9, 11-13, 205-206
  - see also* Ellipsis



**D**

- D-linked 211
  - Aggressively non-d-linked *wh* words 199
- Deletion, of *ne* 141, 146, 155
- Depalatalization 279, 289
- Devoicing 278-279, 287-288, 290, 292-293, 305
- Dialectal variation 36-37, 159, 174-175, 275-294, 304-305, 395-396
- Dissimilation 269
- DP 140, 145-146, 148-150, 152-155, 162

**E**

- Ellipsis
  - CP ellipsis 205-206
    - Island sensitivity of 203
  - Ellipsis condition 212
  - IP ellipsis 197, 357
    - Lack of island sensitivity 206
  - S ellipsis 206-207
- Elsewhere Condition 305-306
- English 17, 25, 45-49, 66 (fn. 3), 69-71, 73-76, 121, 123-130, 132-134, 197, 202-204, 207, 212-214, 361-369, 375
- Epenthesis 17, 104 (fn. 2), 113-115, 264, 276-277, 280-295
  - Epenthetic 104
- EPP 8-12, 122-123
- Event 65, 69, 70-72, 76-77
  - Event decomposition 350-353
  - Event phrase 239, 243, 251-253
  - Event semantics 46, 349-353, 357
  - Event structure 239

- Subevent 46, 239, 245, 251-254
  - see* Event decomposition

## Expletives 4-13

*see also* *Se*

## Extension

*see* Analogy

**F**

- Faithfulness constraints (in OT) 22, 36, 41, 281, 284, 291-293, 307ff.

*see also* Output-Output faithfulness

Resulting from representational entailments 19-20, 23, 26, 29, 39, 43

## Features

- Interpretable 8, 10-11, 13, 245
- Phi 241-242, 245
- Uninterpretable 8-13, 245-246, 252

- Focus 74-75, 210-212, 222-223, 233, 321-330, 337-338
  - Focus condition 212

## Fortition 268

- French 68, 76-79, 160-165, 168-171

French-based creoles 139-140, 145

Middle French 135

Modern French 4-5, 7, 10, 121, 123-136, 139, 158, 225-226, 229

Old French 2-7, 10-13, 135, 149-151, 156, 230, 259, 267

Quebec French 139-141, 146-149

- Futurate 68-69, 72-74

**G**

Galician 299ff.

Galician *geada* 299ff.Galician velarization  
*see* Velarization

Gap

Gascon 109, 111-112, 117-118, 265

Geminate 107, 114

Gemination 24, 27-31, 33, 39,  
41-42Glides 87-88, 96-97, 258, 262, 276,  
281, 290, 292

Greek 218

**H**

Habitual 67-68, 72-74, 77-79

Historical/diachronic linguistics

Historical change 1-14, 108,  
159, 172-174, 395-396Historical phonology,  
morphology 257**I**

Ibero-Romance 220, 222, 224

*Imperfecto* 65-80

Inceptive 69

*see also* Inchoative

Inchoative 65-66, 68-80, 248-250

Infinitival complementation 381-  
382, 385, 391, 394

Inflection 284

Catalan verb inflection 278, 284

Italian verb inflection 18-19,  
21-25, 27-39, 41-43

Rich inflection 10-11

*see also* Paradigm, Conjugation

Intentionality 69, 343-359

Intonation 217-235, 324-325

Inversion, subject 363-364, 367-  
368, 370-376

IP 206-207

*see also* Ellipsis

I principle 129, 134

Italian 17-43, 166, 172, 219, 222-  
223, 234, 245, 247-250

Old Italian 262, 267

*see* Intentionality**J**

Japanese 197, 199-202, 205-206

**K**

Kannada 250-251

**L**Lambda abstraction 364, 370 (fn.  
15)Lateral, palatal 83, 85-89, 91-93,  
96-98, 279, 289

Latin 217-218, 247, 253

Intervocalic nasals 109, 112,  
114 (fn. 8), 115, 117

Left periphery 8-13, 210-212

Lenition 279, 290-292

Lexical

Lexical aspect 242-243, 385-  
386

Lexical frequency 23, 271, 290

Lexical phonology 305-306

Lexical rules 104-106

Lexicon 283

Lexicon Optimization 109, 111,  
282Role in controlling variation 18,  
20, 23-26, 29, 32, 39-41, 43,  
290

Lexico-semantics 45-49

*see* Intentionality

## M

Markedness 284-288, 292, 294

In syllabification 105, 107

Merge 8-9

Cyclic merger 362, 364, 375

Microparameter 159

Micro-parametric variation 139, 159

Misapplication

*see* Underapplication

Morphology 17-26, 42-43, 182-184, 188-189, 194, 275, 284, 287

*see also* Output-Output

faithfulness

*see also* Suppletion

Morphophonemic 178, 184, 191, 195

## N

Nasal 300-302, 307-310

Point of articulation 103-104, 111-113, 117

Navajo 366

Negative

*N*-words 139-150, 156-158

Negative concord 139, 146, 149, 157

Negative polarity items 139, 141-149, 150, 156, 211

## O

Objects 121-123, 129-130

Cognate objects 129-130

Null Objects 123-136

Occitan 226-234, 257, 266

Output-Output

Constraint 113, 116

Correspondence or faithfulness

19-20, 35-39, 41-42, 275, 284, 316-317

*see also* Paradigm uniformity

Stress 35, 41-42

Subsuming morphological relations 20-21, 24, 26

Overapplication 285, 292

Overgeneralization 49-50, 58-62

Undergeneralization 49-50, 58 62

## P

*P* feature 9, 11-12

Palatalization 24-40, 89, 184, 269

Pānini's principle 21-24

Paradigms 25-36, 38, 40, 275-295

Optimal paradigms 276, 284-289, 291-294

Paradigm uniformity 17-21, 35, 39-40, 43, 275, 281, 287, 292

*see also* Analogy

Peak

Alignment 223, 327

Nuclear 324

Preuclear 324

Perception

*see* Dissimilation

*see* Sound change

Portuguese 109, 111-112, 118, 167, 172, 223, 225, 232, 257, 271

Postlexical 104, 106, 187-189

Preterit 41-42, 257, 269

*Preterito* 66-67, 72, 78-80

Velar 257

Proform

Predicative *lo* 355-359

Progressive 67-69, 72-79

- Pronouns  
 Resumptive 376-377  
 Types 242
- Provençal  
*see* Occitan
- Q**
- Quantifier  
 Polyadic 143-144, 146
- Quasi* ambiguity 344-349
- Quechua 322
- Questions 230-234
- R**
- Reconstruction/antireconstruction  
 361-364, 367, 370, 373-374  
*see also* Binding theory
- Recoverability 127
- Relative clause 361-377  
 Correlatives 377  
 Double-headed analysis 361,  
 366, 368-369, 373-374  
 Matching/raising analysis 361,  
 364-366, 368-370, 373
- Remnant 199, 202
- Richness of the Base 35, 109, 282,  
 306
- Romance (languages) 168, 257
- Romanian 167, 171, 177-179, 181,  
 189-190, 194-195, 197-214, 222
- S**
- Sardinian 267
- Se*  
 Doubling of 243-244, 246-248  
 Ergative 247-251
- Expletive 246
- Inchoative  
*see* Ergative
- Reflexive 243-247
- Second language acquisition 49-62
- Sluicing 197-198  
 Lack of island sensitivity 202-  
 203
- Sonority 281, 283, 285-286, 288,  
 291-292
- Sound change 84, 97, 99  
 Perceptually based 290
- Spanish 45-49, 65-80, 160-165,  
 168-171, 224, 232, 234, 241,  
 243-249, 263, 271, 361-362,  
 364, 369-376, 381-396  
 Andean 323  
 Argentine 83-84, 98-99
- Statistical knowledge 21-24, 26,  
 38-40  
 Effects of class size 19, 43  
*see also* Lexical frequency
- Stress 321-339
- Structure preservation 105, 106 (fn.  
 4), 107, 305-306
- Subjects, null 2
- Suppletion (or morphological  
 irregularity) 17-18, 21-23, 25,  
 34-35, 38  
*see also* Pānini's principle
- Syllabification 278, 281-282, 286,  
 295  
 Of intervocalic nasals 105-108
- Syncretism 177-182, 184-185, 189-  
 195  
 Morphological 36-37  
 Paradigm uniformity and 21,  
 32

**T**

Telicity 66-67, 70-80, 245, 248-250, 252  
Topics 133, 210-212  
TP 5, 9, 12  
Transfer 49-51, 58-62  
Transitivity 45-49, 122-123

**U**

Unaccusative/unergative distinction  
45-49  
Underapplication 280, 282-294  
Uniformity  
*see* Paradigm uniformity

**V**

V2 3-5, 7, 9, 13  
Vehicle change 207, 366-367, 374  
Velar 299-318  
Velarization 103-105, 115-117  
Verb raising 9  
Voice phrase 251-253  
Volition 383-384, 387

**W**

*Wh* fronting, multiple 199, 208