

# Language Variation and Contact-Induced Change

## Spanish across space and time

EDITED BY  
Jeremy King  
Sandro Sessarego

JOHN BENJAMINS PUBLISHING COMPANY

# LANGUAGE VARIATION AND CONTACT-INDUCED CHANGE

# CURRENT ISSUES IN LINGUISTIC THEORY

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

ISSN 0304-0763

General Editor

JOSEPH C. SALMONS

University of Wisconsin–Madison

jsalmons@wisc.edu

Founder & General Editor (1975-2015)

E.F.K. KOERNER

Leibniz-Zentrum Allgemeine Sprachwissenschaft, Berlin

efk.koerner@rz.hu-berlin.de

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

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

## *Editorial Board*

Claire Bowern (New Haven, Ct.)

Sheila Embleton (Toronto)

Elly van Gelderen (Tempe, Ariz.)

John E. Joseph (Edinburgh)

Matthew Juge (San Marcos, Tex.)

Martin Maiden (Oxford)

Martha Ratliff (Detroit, Mich.)

E. Wyn Roberts (Vancouver, B.C.)

Klaas Willems (Ghent)

## **Volume 340**

Jeremy King and Sandro Sessarego (eds.)

*Language Variation and Contact-Induced Change. Spanish across space and time*

# LANGUAGE VARIATION AND CONTACT-INDUCED CHANGE

SPANISH ACROSS SPACE AND TIME

*Edited by*

JEREMY KING

*Louisiana State University*

SANDRO SESSAREGO

*University of Texas at Austin / Freiburg Institute for Advanced Studies/*

*Helsinki Collegium for Advanced Studies/*

*Foro Latinoamericano de Antropología del Derecho*

JOHN BENJAMINS PUBLISHING COMPANY  
AMSTERDAM & PHILADELPHIA





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

DOI 10.1075/cilt.340

**Cataloging-in-Publication Data available from Library of Congress:**  
**LCCN 2017046926 (PRINT) / 2017053001 (E-BOOK)**

ISBN 978 90 272 0014 3 (HB)  
ISBN 978 90 272 6455 8 (E-BOOK)

© 2018 – John Benjamins B.V.

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

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

# Table of contents

Introduction	1
<i>Sandro Sessarego &amp; Jeremy King</i>	
<b>Part I. Population migration and contact-induced language change</b>	
CHAPTER 1	
Spatial reconfigurations of Spanish in postmodernity: The relationship to English and minoritized languages	11
<i>Andrew Lynch</i>	
CHAPTER 2	
Female migration and its impact on language choice and use among Afro-Costa Rican women	35
<i>Lisbeth A. Philip</i>	
CHAPTER 3	
Hard come, easy go: Linguistic interfaces in Istanbulite Judeo-Spanish and Afro-Ecuadorian Spanish	63
<i>Rey Romero &amp; Sandro Sessarego</i>	
<b>Part II. Internal and external factors in pragmatic variation</b>	
CHAPTER 4	
Afro-Hispanic contact varieties at the syntax/pragmatics interface: Pro-drop phenomena in Chinchano Spanish	85
<i>Sandro Sessarego &amp; Javier Gutiérrez-Rexach</i>	
CHAPTER 5	
Borrowed Spanish discourse markers in narrative: A comparison across three generations of Tojol-ab'al (Mayan) speakers	111
<i>Mary Jill Brody</i>	
CHAPTER 6	
<i>Hasta perder la última gota de mi sangre: Variation</i> in commissive speech act behavior in Colonial Louisiana Spanish	127
<i>Jeremy King</i>	

## CHAPTER 7

- Discourse markers in variation: *Por tanto* vs. *por lo tanto* 153  
*Sarah Sinnott*

**Part III. Morphosyntactic variation and change**

## CHAPTER 8

- Yo no le conocí a mi abuela*: The use of clitics *le*, *lo*, and *la* in Amazonian  
Colombian Spanish 175  
*Héctor Ramírez-Cruz*

## CHAPTER 9

- Attitudes toward morphosyntactic variation in the Spanish  
of Valencian speakers 199  
*Rosa María Piqueres Gilabert & Matthew Fuss*

## CHAPTER 10

- Stable variation or change in progress? A sociolinguistic  
analysis of *pa(ra)* in the Spanish of Venezuela 223  
*Stephen Fafulas, Manuel Díaz-Campos & Michael Gradoville*

**Part IV. Current issues in bilingual variation**

## CHAPTER 11

- El futuro es perifrástico*: Future tense expression in a bilingual  
U.S. Mexican community 249  
*Kendall Kyzar*

## CHAPTER 12

- L2 Spanish in the U.S. and the question of motivation:  
Changing trends in post-secondary language study 279  
*Terri Schroth & Bryant Smith*

## CHAPTER 13

- Espero estén todos*: The distribution of the null subordinating  
complementizer in two varieties of Spanish 299  
*Adrián Rodríguez Riccelli*

- Index 335

# Introduction

Sandro Sessarego<sup>1,2,3,4</sup> & Jeremy King<sup>5</sup>

<sup>1</sup>University of Texas at Austin / <sup>2</sup>Freiburg Institute for Advanced Studies /

<sup>3</sup>Helsinki Collegium for Advanced Studies / <sup>4</sup>Foro Latinoamericano de

Antropología del Derecho / <sup>5</sup>Louisiana State University

The present volume is a collection of original contributions dealing with Hispanic contact linguistics. This work seeks to further develop the intense interest over recent decades in aspects of Spanish philology related to contact and variationist linguistics, and it not only provides a current perspective on the status of Spanish linguistics and its significant contributions to sociolinguistic, contact and applied linguistic theories of language variation and change but also presents new perspectives on a number of issues that have long been at the core of (socio)linguistic research: language contact, morphosyntactic and pragmatic variation, bilingualism, second language acquisition and language change.

The papers included in this volume emerged from presentations and discussions held at the *XXIX Biennial Louisiana Conference on Hispanic Languages and Literatures*, February 27–March 1, 2014, hosted by Louisiana State University. The conversations that took place at this meeting underscored the dominant role that the themes of language variation and change have taken on within numerous subfields of linguistics in the early years of the 21st century. The motif of the role of internal and external factors in the variation evident in the Spanish language emanated from the sessions on Hispanic linguistics at this meeting and laid the foundation for this collaborative publication.

Although the notion of variation in the Spanish language has been well documented over the past century (Alonso 1935; Kany 1945; Penny 2000; Zamora Vicente 1960; among many others), much of the linguistics literature has come to be dominated by this theme, to the extent that it is now thought of by many specialists as a subfield of the discipline in its own right. The relatively recent appearance of journals such as *Language Variation and Change* and *Linguistic Variation* attest to the fact that variationism has become a major concern of cross-linguistic scholarship of late. Innumerable publications have addressed various aspects of this discipline and have widened the academic community's breadth of knowledge related to Hispanic linguistics and in particular the myriad varieties of Spanish that exist in the world. Among these we find much noteworthy work on variation within the Spanish-speaking world (Anipa 2001; Lipski 2008; Penny 2000), volumes

on contact varieties of Spanish (Klee & Lynch 2009; Otheguy & Zentella 2012; Roca & Lipski 1993; Sessarego & González-Rivera 2015; Silva-Corvalán 1995) and studies on evolution and change within the language (Clements 2009; Company 1997; Mar-Molinero 2000).

Variationist linguistics seeks to address a variety of exigent, and interrelated, issues regarding the nature of language; it is our contention that interdisciplinary approaches are best suited to answer these questions. The following guided the development of the current volume:

- How does language change come about in a given linguistic community?
- What is the relative importance of internal and external factors in the types of language change that are manifested in Spanish?
- How is language variation viewed by speakers of different communities in the Spanish-speaking world?

This volume seeks to provide insight into these issues through the use of a variety of approaches. A number of studies included here employ established data collection methods, including anthropological/ethnographic techniques (Brody; Philip), sociolinguistic interviews (Fafulas, Díaz-Campos & Gradoville; Kyzar; Romero & Sessarego), collection of written corpora (King) and surveys (Schroth & Smith; Piqueres Gilabert & Fuss). The linguistic elements chosen for analysis in several chapters are familiar ones in the current literature on Spanish sociolinguistics: subject and object pronouns (Sessarego & Gutiérrez-Rexach; Ramírez-Cruz; Romero & Sessarego), future verb tense (Kyzar), speech acts (King) and discourse markers (Sinnott; Brody). In addition to original case studies offering new takes on well-known issues, several studies in this volume introduce novel data collection methods, topics and/or theoretical stances. Rodríguez Riccelli and Sinnott analyze different elements of language with corpora gathered via social media. Romero & Sessarego and Sessarego & Gutiérrez-Rexach combine traditional sociolinguistic methodologies with formalist approaches to variation. Piqueres Gilabert & Fuss introduce a thought-provoking twist on the traditional matched-guise technique. And Lynch mixes sociolinguistic approaches with critical theory in his exploration of modern Spanish.

The volume comprises four sections, organized according to the type of phenomena under analysis: (I) Population Migration and Contact-Induced Language Change; (II) Internal and External Factors in Pragmatic Variation; (III) Morpho-syntactic Variation and Change; and (IV) Current Issues in Bilingual Variation. Though the topics of many papers included overlap with the themes of more than one of these sections, they have been organized with regard to the most salient issue discussed by their author(s). The volume covers an array of Spanish varieties

distributed across North, South and Central America; the Caribbean; the Iberian Peninsula and the Bosphorus. It deals with both native and non-native varieties of the language and includes both synchronic and diachronic studies. The investigations reported on in the present collection evince a cohesiveness of thought while tackling distinct research questions of current interest within the field. The volume addresses, and challenges, current theoretical assumptions on the nature of language variation and contact-induced change through empirically-based linguistic research.

Andrew Lynch opens Section I by offering his view on the nature of Spanish spatial reconfigurations in relation to English and minoritized languages in the global era, when geopolitical boundaries and national identities are becoming more blurry and, consequently, less easily distinguishable. The author claims that the advent of information and communication technologies and the diffusion of mass media and the internet are producing a change in the way societies perceive languages and use them. This has an inevitable effect on the way Spanish coexists with English and other languages spoken across Spain and Latin America, such as Catalan and Quechua. The new scenarios emerging from this process, oftentimes, do not follow normative standards (Gee 2012); rather, they generate patterns of variation that appear to be solved by the constant contrast of local and global contexts, thus resulting in several 'glocal' realities.

In Chapter 2, Lisbeth Philip analyzes patterns of language maintenance and language shift (LMLS) among the female members of the Afro-Costa Rican immigrant community, whose parents or grandparents moved from the West Indies to Siquirres and Puerto Limón, Costa Rica. The author considers a variety of social factors such as family origins, education and language proficiency to understand to what extent English and Spanish coexist in these communities. Findings suggest – in line with Fishman (1972) – that, in order to understand the dynamics of LMLS in a given community, it is crucial to identify and separate the linguistic domains that are relevant to such a social group. In particular, the author suggests that the family domain and the affective domain (including close friendship, church, etc.) appear to be the most conservative ones in terms of English language maintenance and highly depend on the language that speakers studied when they attended school. These claims are also in line with Duncan's (2005) observations on the importance of family, religion and education in determining Afro-Costa Rican culture and identity.

Rey Romero and Sandro Sessarego, in Chapter 3, analyze subject pronoun use and variable phi-agreement phenomena in Afro-Ecuadorian Spanish (AES) and Istanbulite Judeo-Spanish (IJS). The authors account for the similarities found in these two varieties, which have never been in contact with each other, by proposing a general model of cross-generational contact-induced language change

(Sessarego 2013), which relies on current assumptions on the architecture of the faculty of language (Chomsky 1995), its modularity (Reinhart 2006) and the vulnerability of modular interfaces (White 2009). They conclude that the features found in AES and IJS can be conceived as nativized traces of advanced second language acquisition strategies and early first language acquisition processes.

The second section of the volume deals with pragmatic variation, which begins with Sessarego and Gutiérrez-Rexach's study. The authors adopt the general contact-induced transmission model proposed in the previous chapter to account for a specific configuration of pro-drop phenomena found in Chinchano Spanish (CS), an Afro-Hispanic dialect spoken in Chinchá, Peru. In particular, this work suggests that, at a certain point in the evolution of CS, there must have been a generation of Spanish second language speakers who did not completely achieve a target-like proficiency in the use of null and overt subjects – since it implies a heavy processing cost at the syntax-pragmatics interface (Sorace 2003). The linguistic output of such speakers served as the primary linguistic data for the following generation of Afro-Chinchanos, who learned this contact variety as a native language. This process resulted in the instantiation of a grammar showing properties that do not align this dialect with either null-subject languages (NSLs) like Spanish or non-null-subject languages (NNSLs) like English (Rizzi 1982).

In Chapter 5, Mary Jill Brody focuses on the use of borrowed Spanish discourse markers (DMs) in the speech of three generations of Tojolab'al (Mayan) speakers. She discovers that, even though the youngest generation presents overall higher levels of borrowed words, the discourse structure of Tojolab'al and the use of borrowed DMs has remained substantially constant across the three groups analyzed. The presence of all five major functions of borrowed DMs identified by Fraser (2006) and the author in all three generations of Tojolab'al speakers, as well as a significant amount of overlap in the employment of Spanish DMs such as *entonces*, *pues* and *porque*, support Brody's claim that the basic discourse structure of Tojolab'al has remained stable across the span of three generations.

Chapter 6, by Jeremy King, presents an historical accounting of one type of speech act, the commissive, in Colonial Louisiana Spanish. Based on a corpus of business letters written by government officials, King argues that the structure and nature of the commissives issued during this time period were in large part determined by the power relationship between the letter writer and the addressee. Employing the terminology of Bilbow (2002), the author illustrates that commissives issued by inferiors nearly always took the form of *offers* (speech acts formulated using direct linguistic encoding but uninitiated by a prior speech act), while superiors tended to issue *promises* (direct speech acts in which the speaker responds to an addressee's prior request). Additionally, the author notes that inferiors customarily committed themselves to a

course of action using the morphological future tense, while superiors more frequently employed the present tense or indirect formulations to express self-commitment.

In the following chapter, Sarah Sinnott presents an analysis of the Spanish DMs *por tanto* (PT) and *por lo tanto* (PLT). She shows that, even though traditional grammars tend to classify these constructions as synonymous (cf. Martín Zorraquino & Portolés Lázaro 1999), a closer look at their actual use in context reveals a number of functional differences between them. Specifically, Sinnott claims that these two variants differ in terms of the type of causality expressed in the utterances in which they are used (Mak & Sanders 2013). In fact, PLT is used significantly more in cases of epistemic causality and tend to contribute higher argumentative strength than PT.

Section III focuses on questions of morphosyntactic variation. This section begins with Héctor Ramírez-Cruz's chapter, which provides a quantitative analysis of the use of third-person pronominal objects (*lo*, *la*, *le*) in Amazonian Colombian Spanish. The author analyzes the speech of two different communities in the Colombian Department of Amazonas (the villages surrounding Puerto Nariño and Leticia) and codes his sociolinguistic data according to a variety of internal and external factors. After providing a series of statistics aggregated via R-brul, he concludes that *léismo* and *loísmo* are the most common phenomena found across the communities investigated. He hypothesizes that such patterns may have originally been introduced in the region during the rubber boom of the late 19th and 20th centuries by speakers coming from Peru and other regions of Colombia, who settled there to work in the rubber industry. The alternation and selection among *lo*, *la* and *le* is analyzed in light of Mufwene's (2001) model of language evolution via feature competition.

Chapter 9 is an analysis of the attitudes toward morphosyntactic variation in the Spanish of Valencian speakers. In this work, Rosa Piqueres Gilbert and Matthew Fuss study three specific variables: use of the locative prepositions *a/en* "at/in", partitive *de* "of" and expletive *que* "that" in interrogative sentences. They adopt a matched-guise technique (Lambert et al. 1960) to evaluate the attitudes of 31 Valencian-Spanish bilinguals from La Plana, Castellón. Results suggest that speakers' judgments do not appear to be significantly affected by either the standard or the non-standard variants of the partitive *de* and the relative *que* variables, while the use of either *a* or *en* as locative prepositions is perceived as a strong indicator of socioeconomic class.

In Chapter 10, Stephen Fafulas, Manuel Díaz-Campos and Michael Gradoville provide a quantitative analysis of *pa/para* "for" alternation in Caracas Spanish. After providing a series of statistical runs using generalized linear mixed-effects models in R, they conclude that the alternation between these two variants is



quite stable in the Venezuelan speech community analyzed. In fact, this linguistic variable appears to be regularly stratified according to social class and speech style and does not seem to be symptomatic of a linguistic change in progress.

The last section of the volume is dedicated to variation in bilingual speech. In Chapter 11, Kendall Kyzar provides a Goldvarb analysis of variable future verb forms in the speech of the Mexican community living in Baton Rouge, Louisiana. The author considers a number of internal and external factors and concludes that the periphrastic future form is gaining terrain at the expense of morphological future form. After comparing the stage of this phenomenon with a number of other studies on futurity (Gutiérrez 1995; Orozco 2004, 2007, etc.), he suggests that the Baton Rouge case may be conceived as a change in progress, which is likely already occurring in monolingual Mexican communities, and appears to be accelerated by the contact with English among the Mexicans living in this Louisiana city.

Terri Schroth and Bryant Smith, in the following chapter, analyze the motivations of American students in studying Spanish at the university level. They compare Spanish students on two campuses in different parts of the country. After analyzing the results of an open-ended questionnaire, they find that both *integrative* and *instrumental* motivations form part of students' interest in learning Spanish in the 21st century. However, they note that contemporary students are much more likely to indicate instrumental reasons for studying the language (such as concerns regarding their careers) than their peers in similar studies from the latter half of the 20th century (cf. McEwan & Minkle 1979). These results thus suggest a transitioning perspective on the role of Spanish in the U.S. in the modern era.

Finally, in Chapter 13, Adrián Rodríguez-Riccelli analyzes and compares the patterns of *que*-drop variation in two Spanish speaking communities: the English-Spanish bilingual community of Los Angeles and the monolingual Spanish community of Mexico City. He relies on Twitter to create two comparable corpora of data and on R-brul to study the internal and external factors regulating the alternation between overt and covert *que* in these two dialects. The author shows that both *que* and  $\emptyset$  are competing variants in the grammars of Los Angeles Spanish and Mexico City Spanish which appear to be conditioned by the same syntactic and lexical factors, such as modality of the matrix verb and the status of the embedded subject. Contrary to expectations (cf. Silva-Corvalán 1994), Rodríguez-Riccelli finds that participants' city of residence does not have any significant correlation with this linguistic phenomenon, thus suggesting that contact with English is not necessarily increasing the probabilities of *que*-dropping in Spanish. On the other hand, what contact may be held responsible for is the increase in the range of contexts in which *que*-drop may occur, thus suggesting that a generalization of the favorable context for  $\emptyset$  may be underway in Los Angeles Spanish.

The sustained contact between Spanish and other languages in different parts of the world has given rise to a wide variety of changes in the language. The types of variation evinced by the studies in this collection are the result of many different factors and processes. With the research reported on here, we hope to provide new insight into a variety of issues involving contact and change in the Spanish-speaking world.

## References

- Alonso, Amado. 1935. *El problema de la lengua en América*. Madrid: Espasa-Calpe.
- Anipa, Kormi. 2001. *A Critical Examination of Linguistic Variation in Golden-Age Spanish*. (Berkeley Insights in Linguistics and Semiotics 47). New York: Peter Lang.
- Bilbow, Grahame T. 2002. "Commissive Speech Act Use in Intercultural Business Meetings". *International Review of Applied Linguistics* 40.287–303. doi:10.1515/iral.2002.014
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Clements, J. Clancy. 2009. *The Linguistic Legacy of Spanish and Portuguese: Colonial Expansion and Language Change* (Cambridge Approaches to Language Contact). Cambridge: Cambridge University Press. doi:10.1017/CBO9780511576171
- Company, Concepción, ed. 1997. *Cambios diacrónicos en el español*. Mexico City: Universidad Nacional Autónoma de México.
- Duncan, Quince. 2005. "El negro antillano: inmigración y presencia". *El negro en Costa Rica* ed. by Quince Duncan & Carlos Meléndez, 11th ed., 115–175. San José: Editorial Costa Rica.
- Fishman, Joshua A. 1972. *Language in Sociocultural Change*. Stanford, Calif.: Stanford University Press.
- Fraser, Bruce. 2006. "Towards a Theory of Discourse Markers". *Approaches to Discourse Particles* ed. by Kerstin Fischer, 189–204. Amsterdam: Elsevier.
- Gee, James Paul. 2012. *Social Linguistics and Literacies: Ideology in Discourses*, 4th ed. London: Routledge.
- Gutiérrez, Manuel J. 1995. "On the Future of the Future Tense in the Spanish of the Southwest". *Spanish in Four Continents: Studies in Language Contact and Bilingualism* ed. by Carmen Silva-Corvalán, 214–226. Washington: Georgetown University Press.
- Kany, Charles E. 1945. *American Spanish Syntax*. Chicago: University of Chicago Press.
- Klee, Carol A. & Andrew Lynch. 2009. *El español en contacto con otras lenguas*. Washington: Georgetown University Press.
- Lambert, Wallace, Richard C. Hodgson, Robert C. Gardner, & Samuel Fillenbaum. 1960. "Evaluational Reactions to Spoken Languages". *Journal of Abnormal and Social Psychology* 60:1.44–51. doi:10.1037/h0044430
- Lipski, John M. 2008. *Varieties of Spanish in the United States*. Washington: Georgetown University Press.
- Mak, Willem M. & Ted J.M. Sanders. 2013. "The Role of Causality in Discourse Processing: Effects of Expectation and Coherence Relations". *Language & Cognitive Processes* 28:9.1414–1437. doi:10.1080/01690965.2012.708423
- Mar-Molinero, Clare. 2000. *The Politics of Language in the Spanish-Speaking World: From Colonisation to Globalisation*. London: Routledge.

- Martín Zorraquino, María Antonia & José Portolés Lázaro. 1999. "Los marcadores del discurso". *Gramática descriptiva de la lengua española* ed. by Ignacio Bosque & Violeta Demonte, vol. III, 4051–4213. Madrid: Espasa-Calpe.
- McEwan, Angela & Roger Minkle. 1979. "Why am I Learning Spanish? An Evaluation of Student Response." *Hispania* 61.89–97. doi:10.2307/340083
- Mufwene, Salikoko S. 2001. *The Ecology of Language Evolution*. London: Cambridge University Press. doi:10.1017/CBO9780511612862
- Orozco, Rafael. 2004. "A Sociolinguistic Study of Colombian Spanish in Colombia and New York City". Ph.D. dissertation, New York University.
- Orozco, Rafael. 2007. "The Impact of Linguistic Constraints on the Expression of Futurity in the Spanish of New York Colombians". *Spanish in Contact: Policy, Social, and Linguistic Inquiries* ed. by Richard Cameron & Kim Potowski, 311–328. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/impact.22.22oro
- Otheguy, Ricardo & Ana Celia Zentella. 2012. *Spanish in New York: Language Contact, Dialectal Leveling, and Structural Continuity*. (Oxford Studies in Sociolinguistics). Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199737406.001.0001
- Penny, Ralph. 2000. *Variation and Change in Spanish*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139164566
- Reinhart, Tanya. 2006. *Interface Strategies: Reference-Set Computation*. Cambridge, Mass.: MIT Press.
- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris. doi:10.1515/9783110883718
- Roca, Ana & John M. Lipski, eds. 1993. *Spanish in the United States: Linguistic Contact and Diversity* (Studies in Anthropological Linguistics 6). Berlin & New York: Mouton de Gruyter. doi:10.1515/9783110804973
- Sessarego, Sandro. 2013. *Chota Valley Spanish*. Madrid & Frankfurt: Iberoamericana/Vervuert.
- Sessarego, Sandro & Melvin González-Rivera, eds. 2015. *New Perspectives on Hispanic Contact Linguistics in the Americas*. (Lengua y sociedad en el mundo hispánico 35). Madrid & Frankfurt: Iberoamericana/Vervuert.
- Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. New York: Oxford University Press.
- Silva-Corvalán, Carmen, ed. 1995. *Spanish in Four Continents: Studies in Language Contact and Bilingualism*. Washington: Georgetown University Press.
- Sorace, Antonella. 2003. "Near-Nativeness". *The Handbook of Second Language Acquisition*, ed. by Catherine J. Doughty & Michael H. Long, 130–153. Oxford: Blackwell. doi:10.1002/9780470756492.ch6
- White, Lydia. 2009. "Grammatical Theory: Interfaces and L2 Knowledge". *The New Handbook of Second Language Acquisition* ed. by William C. Ritchie & Tej K. Bhatia, 49–68. Bingley, UK: Emerald Group.
- Zamora Vicente, Alonso. 1960. *Dialectología española*. Madrid: Gredos.

PART I

# Population migration and contact-induced language change



## CHAPTER 1

# Spatial reconfigurations of Spanish in postmodernity

## The relationship to English and minoritized languages

Andrew Lynch

University of Miami

In this chapter, I propose four principal ways in which the Spanish language is being spatially reconfigured vis-à-vis English and local minoritized languages in postmodernity: (1) the ideological relocation of Spanish in the global ‘marketplace’ economy; (2) the propagation of English-influenced discursive patterns within the context of late capitalism, e.g., preference for familiar forms of pronominal address (*tuteo*) in commerce; (3) the incursion of English in public life and education, placing traditionally minoritized languages in sharper relief in the linguistic landscape; and (4) the normativization and normalization of Spanish in cyberspace and in global mass media, where English is hegemonic. I suggest that these ideological phenomena will condition structural and discursive variations in Spanish in the years ahead.

### 1. Introduction

In the postmodern *qua* global era, the ideological spaces that language occupies begin to be reworked or reconstructed. As late capitalist and transnational orders of things slowly emerge, the idea of ‘language’ in the 21st century gradually becomes more dependent upon economic orders of things (i.e., relegated to the ‘market’) and less easy to ‘locate’ in terms of geopolitical boundaries or criteria of nationality and nationhood. Throughout the Spanish-speaking world, post-modern relocations of language are not only creating reconfigurations of Spanish vis-à-vis local minoritized languages but are also concomitant with incursions of English in the discursive arenas of Spanish, setting the stage for wide-scale English-language contact phenomena in the decades ahead. These novel spaces

of language contact are multiple and, in cultural and social terms, highly powerful: the late capitalist marketplace, the booming service and tourism sectors of the global economy, the educational realm, the Internet and mass media. Although the growing presence of English further displaces traditionally minoritized languages, at the same time it provides an ideological arena for resistance and the reworking of identities, as bilingual/bicultural repertoires are placed in sharper relief and relocated in some ways beyond the 'locales' previously drawn by modernity.

My objective is to paint a broad picture of the shifting ideological construction and locations of Spanish and its relationship to English and minoritized languages. The theoretical crux of my argument is that the late capitalist 'global marketplace' economy and the advent of information and communication technologies, cyberspace and the proliferation of mass-mediated popular culture have begun to produce a shift of ideological paradigm in the ways that we conceptualize language culturally and perceive and imagine the use of language (cf. Appadurai 1996; Heller & Duchêne 2012). The chapter is divided into four sections that pose what I have identified as major facets of the Spanish language in postmodernity: (1) the spatial relocation of Spanish vis-à-vis English in the discourse of the global(ized) economy, as made evident in the ideological shift from language as personal possession to marketplace commodity in the U.S. context; (2) the propagation of English-influenced patterns of Spanish language use due to the pervasiveness of English in postmodern spaces, illustrated by the rapid expansion of familiar forms of pronominal address (*tuteo*) in the discourse of late capitalism throughout much of the Spanish-speaking world; (3) the incursion of English in public life and education throughout the Hispanic world, prompting the rearticulation of space for minoritized languages such as Catalan and Quechua; and finally, (4) the problematic normativization and normalization of Spanish in cyberspace and in global mass media, where English is hegemonic.

## 2. The relocation of Spanish in postmodernity: From nation to market

Simply stated, the existence of a language as something real or material, i.e., language as object, is predicated upon the act of establishing a spatial location for it. If we believe that Spanish exists, it must exist SOMEWHERE. One might argue that Julióbreaga was the first such 'somewhere' for the Spanish language. Julióbreaga was one of the central anchor points for the Roman presence in the north of the Iberian Peninsula and, according to some historians, the symbolic site of the Romans' final struggle to subdue the last of the indigenous resistance in Hispania, the

Cantabrians. It was the physical site of Roman incursion into the arena of those fierce warriors viewed as the cultural and linguistic Other. This same place would later be reputed as “la cuna del castellano” (the cradle of Castilian), and by 1492 Antonio de Nebrija would famously affirm that “la lengua fue siempre compañera del Imperio” (language always accompanied empire). Just a few years later, the *conquistadores* sent by the Crown of Castile would erect a cathedral over the destroyed bases of the temple of Tenochtitlán, today the center of Mexico City, as a new order of things gradually emerged in Western Europe. Language, in its written form, began to provide a kind of blueprint for the terrestrial shape of modern nation-states, geopolitically defined, imagined as bounded and sovereign communities onto which market-style readerships were mapped (Anderson 1983). These readerships, modeled on local vernacular varieties, provided the basis for print capitalism and concomitantly the basis for nations as they would be imagined in the age of modernity. As Anderson affirmed: “Nothing served to ‘assemble’ related vernaculars more than capitalism, which, within the limits imposed by grammars and syntaxes, created mechanically reproduced print-languages capable of dissemination through the market” (1983: 40).

The printed word, now poised to extend its reach beyond a Latin-literate elite, established a mooring for languages as they would be imagined, conceptualized, normativized and contemplated in geopolitics throughout the 19th and 20th centuries.<sup>1</sup>

In the 21st-century logic of late capitalism, the metaphor of ‘market’ creates a spatial representation for language, one that is somehow less spiritual or transcendental than 19th and 20th-century spaces of modern nationhood (cf. Vossler 2010). Economic markets are conceived in rather more concrete terms: as numerical and statistical realities, in which physical or symbolic goods are bought, sold and circulated. *Telefónica’s* decision to provide customer services in Catalan following a local government call for bids by the *Generalitat de Catalunya* in 1998 is a good example. The privatization of Spain’s national telephone services meant liberalization of restrictive centralist policies, whereby “the symbolic value of Spanish as a national resource and as resource for social identification and differentiation become secondary” to consumer market considerations (Pujolar 2007: 75). Da

---

1. Roman Jakobson elaborated upon the ideological basis of the phoneme: the bounded and sovereign sound as imagined or made manifest in the mind of the speaker who belongs to a homogeneous speech community. This concept provided a framework for structuralist analysis in linguistics during the 20th century and in other disciplines such as anthropology, sociology, psychology, literary studies and art criticism.



Silva et al. (2007) made a similar observation regarding the situation of French in Canada:

In an economy which increasingly emphasizes global customer service, we witness the commodification of language and its disentanglement from ‘personhood’ ... . In the new economy, linguistic practices can be seen as resources which, in an effort to cater to markets and consumers, are mobilized by actors to produce and reproduce a multiplicity of social identities. (185)

This ‘disentanglement from personhood’ is one aspect of the relocation of language that differentiates modernity from postmodernity. The statement that ‘there is a market demand for Spanish in the US economy’ is likely received by most as a numerical reality that has rather little to do with personal belief or ethnic identity. It is simply perceived as a ‘fact’ of the market: people who speak Spanish have the power of a dollar and the power of a vote. This is not the way that Spanish was thought of only three decades ago.

The situation of Spanish in Miami is highly illustrative.<sup>2</sup> A keyword search of articles published in *The Miami Herald* reveals a dramatic decrease in the number of articles containing the terms ‘Hispanic’, ‘English’ and ‘assimilation’ from the late 1980s to the decade of the 1990s (see Figure 1). In the late 1980s, more than thirty such articles appeared; during the early 1990s, only nine appeared.

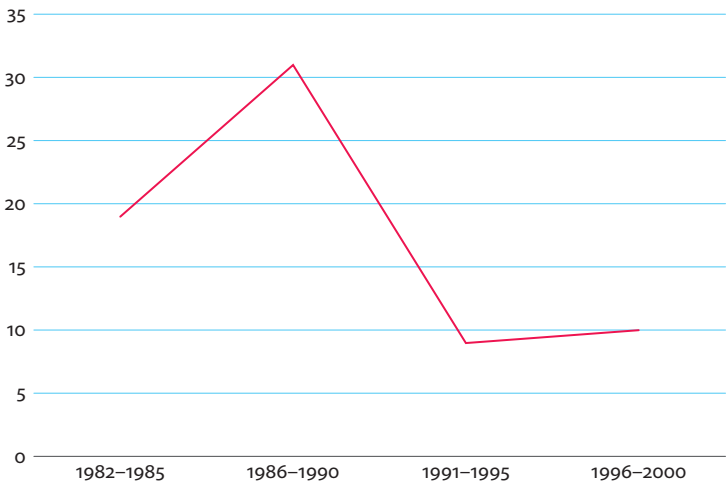


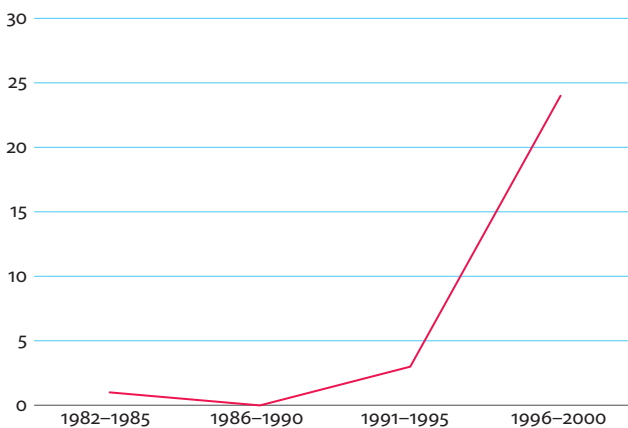
Figure 1. Miami Herald articles containing the terms ‘Hispanic’, ‘English’, ‘assimilation’

2. See Carter & Lynch (2015) for a recent overview of the sociolinguistic situation of multilingual Miami.

In the story headlines seen in the late 1980s, there is a clear ideological tendency to emphasize assimilation to an Anglo cultural mainstream and concomitantly to devalue Spanish language use in Miami, as in the following examples:

- a. “Ferre tells editors’ group Hispanics must assimilate”
- b. “Power of the ‘melting olla’ works for Hispanics, too”
- c. “Latins begin to mingle socially with Anglos”
- d. “Second generation of Cuban exiles more ‘Miami-ized’”
- e. “If you are born with Spanish, it’s no special deal”

During the late 1980s, the same period in which articles containing the terms ‘Hispanic’, ‘English’ and ‘assimilation’ increased dramatically (Figure 1), there were no articles published containing all of the terms ‘Spanish’, ‘bilingual’, ‘education’ and ‘economic’ (see Figure 2). In the early 1980s, only one such article appeared. Dramatically, however, by the late 1990s, 24 articles were published containing those terms.



**Figure 2.** Miami Herald articles containing the terms ‘Spanish’, ‘bilingual’, ‘education’, ‘economic’

There is an obvious shift of attention away from the socially controversial issues of Anglophone assimilation posed during the 1980s to the urgent economic importance of Spanish in South Florida’s burgeoning global economy by the late 1990s (cf. Lynch 2000). In less than a decade, the ideological construction of Spanish in Miami transformed from a cultural liability of sorts into a valuable asset in the city’s job market and the region’s economic infrastructure (cf. Smith & Schroth, this volume). The following excerpt from a 1997 article reflects this phenomenon quite clearly:

Depending on who's doing the talking, language is either an essential qualification, a cause to be fought for, or a complicated problem that managers are finally learning to handle. ... Most important: Bilingualism is a big asset. ... In some professions, speaking only one language can prevent you from moving ahead.

(7 September 1997)

The headline of this article – “Bilingualism a huge asset in job search” – stands in stark ideological contrast to the 1980s headline cited above (e): “If you are born with Spanish, it's no special deal”.

As South Florida's economy grew increasingly linked to the economies of diverse Latin American countries during the 1990s, the Spanish language became objectified as a commodity on a global market (cf. Del Valle 2007). This ideology was already evident by 1991, as reflected in the following excerpt:

Spanish classes, in fact, are a thriving business in South Florida. ... With domestic business mired in economic doldrums, American companies have had to look overseas for new markets, and in Miami that means Central and South America. For that reason, many corporate budget slashers have gone easy on line items for language training.

(18 July 1991)

Clearly, the cultural concern that Spanish would not go away during the 1980s turned into an economic worry that Spanish would not be present enough by the late 1990s. A 1999 story headline stated it precisely as such: “Business leaders worry that [Spanish] language skills aren't at suitable level” (27 June 1999). An article published three years earlier highlighted the economic risk for the region if bilingualism were not sustained in South Florida: “Dade County's future economic growth as an international marketplace could be in jeopardy if young Hispanic Americans continue to grow up with insufficient Spanish-language skills” (19 March 1996).

In sum, there is clearly a change in the way that Spanish is ideologically constructed and socially perceived in public life in South Florida from the late 1980s to the late 1990s, i.e., the language becomes less of a talking point regarding ethnic identity and cultural assimilation to become enmeshed in the discourse of the global economy. The principal metaphor shifts from nationalism to that of capitalism, from the region as a ‘community’ to the region as a ‘marketplace’. Marketplace ideologies and the crucial role of language in business and commercial spheres were not novel in the 1990s, however. As soon as the first Cuban exiles arrived in Miami in the early 1960s, a discourse of solidarity in Spanish quickly coalesced (Lynch 2009; Laguna 2010). In my view, one of the key features of that discourse was *tuteo*, or the preference for *tú* forms of address over *usted*, fueled by the sales-centered ideology of capitalism. Indeed, I would argue that Cuban Spanish *tuteo* in 1960s–1970s Miami provided a glimpse into a

more general discursive pattern to be observed throughout much of the Spanish-speaking world, including other communities in the U.S., by the arrival of the new millennium.

### 3. The discursive effects of late capitalism: *tuteo*

In their seminal argument regarding the expansion of *T* (i.e., familiar) forms of address in Western European languages,<sup>3</sup> Brown & Gilman (1960) maintained that:

In a fluid society crises of address will occur more frequently than in a static society, and so the pronominal coding of power differences is more likely to be felt as onerous ... . A strong equalitarian ideology of the sort dominant in America works to suppress every conventional expression of power asymmetry. (268)

Numerous studies have suggested that indices of upward social mobility correlate with a preference for the use of *tú* forms; *usted* becomes more discursively marked as upward social mobility or the possibility of upward social mobility is attained (Fernández 2003). The contemporary history of Spain is a case in point. Probably the most *tuteísta* country in the Spanish-speaking world in the present day – and economically the most prosperous from the 1980s until the late 2000s – Spain remained economically stagnant under Francisco Franco's totalitarian regime (1939–1975). Until the democratic transition of the late 1970s, *usted* was widely used, especially in rural areas. The introduction of a capitalistic economy and social democracy spurred a linguistic change greatly favoring *tuteo* among young and middle-aged people, in major cities and in the commercial spheres of public life, creating an abrupt generational shift marked by a rural/urban divide.<sup>4</sup>

The possibility of upward social mobility offered by urban areas is, in all likelihood, the reason why Spanish-speaking urbanites tend to be more *tuteísta* than their rural counterparts. Pronominal usage in Bogotá, Mexico City and San José, Costa Rica reflects this tendency even in countries where *usted* forms tend to

---

3. Brown & Gilman (1960) contrasted 'familiar' (*T*) pronoun usage with 'polite' (*V*) forms of address in French (*tu* vs. *vous*), German (*du* vs. *Sie*) and Italian (*tu* vs. *Lei*). In Spanish, the corresponding forms are *tú* (*T*) and *usted* (*V*); the practice of addressing an interlocutor as *tú* in Spanish is termed *tuteo*.

4. This shift is saliently portrayed in the dialogues of the long-running popular Spanish television series *Cuéntame cómo pasó* (TVE), which recounts the turbulent post-Franco years as experienced by a family in the country's capital and largest city.

prevail. There is a “fiction of a friendship” (Valdés & Pino 1981) inherent in the late capitalist US model of marketing, advertising and customer service (see Figure 3). I argue that this ideology foment the expansion of *tuteo*.



Figure 3. Banco de Panamá's 'Caja amiga'

Rodríguez Solórzano (1999) and Murillo Medrano (2003) both comment on the incursion of *tú* in advertising and commerce in the Costa Rican capital, where *vos* and *usted* have traditionally predominated. Rodríguez Solórzano (1999) urges Costa Rica's principal newspaper *La Nación* to stop confusing children and banish *tuteo* from the country in its advertisements: “¿Por qué *La Nación* no hace un esfuerzo por desterrar el tuteo en su publicidad?” (Why doesn't *La Nación* make an effort to get rid of *tuteo* in its advertising?). He goes on to berate several product industries, Coca Cola among them, for using *tú* command forms rather than the more intimate and affectionate *vos* or polite *usted*. In commercial service encounters in Costa Rica, Murillo Medrano observes that “the use of *tú* ... in speech situations such as cashier-customer or salesperson-client seems to indicate that there is a realm in which the other two pronouns do not meet the requirements of face work carried out by the speakers involved” (2003: 136).<sup>5</sup> Murillo Medrano proposes that, in the Costa Rican context, *tú* emerges as the form that permits closeness without getting too close (*vos*) or remaining too distant (*usted*). In Mexico, Cuba and other Latin American countries where *vos* is never used, *tú* would clearly be the only viable option for breaching the distance

5. “El empleo de *tú* ... en situaciones de habla del tipo cajero-cliente o dependiente-cliente parece indicar que existe una esfera en la que los otros dos pronombres no cumplen con los requisitos de las actividades de imagen [face work] que los hablantes llevan a cabo”.

implicated in doing business with strangers and establishing an imaginary solidarity or feigned familiarity for the sake of gaining the confidence of a customer or client (cf. King, this volume).

In a comparative speech act analysis of compliment responses in spontaneous conversation, Valdés & Pino observed in 1981 that US Mexican-American bilinguals (in the Mesilla Valley of southern New Mexico) seemed much less concerned about marking social distance than did their monolingual Mexican counterparts (in Ciudad Juárez, located just across the Mexican border). They affirmed that in their usage of pronominal address, the US bilinguals “carry out a fiction that involves treating others as if they were close friends” (62). Valdés & Pino argued that, if Americans indeed tend to conform to the ‘be friendly’ politeness rule proposed by Lakoff (1973), then their interactions “will reflect the use of strategies that seek to create an impression of camaraderie with most interlocutors, rather than one of social distance” (58). This general interactional tendency stands in contrast to that of Mexican society, according to Valdés & Pino, which they claim is “more highly stratified” and thus may reflect more general adherence to Lakoff’s (1973) first rule of politeness: ‘don’t impose’ (58). Hardin (2001) reached a very similar conclusion in a comparative analysis of the discourse strategies of persuasion observed in Spanish-language television advertising in the US, Spain and Chile. She concluded that “US and Spanish advertisers consider egalitarianism to be highly valued by their viewers. In contrast, the higher degree of negative politeness, power, and distancing strategies in Chile may reveal a society seen by advertisers as being stratified and class conscious” (180). In Chile, as in the rest of the nations of the Spanish-speaking world, knowledge and use of English have become in recent years a *sine qua non* for participation in the global economy, particularly for the urban middle and upper classes. I turn now to the expansion of English as a fundamental characteristic of ‘progress’ in Latin America, an imposition – also possibly read as privilege – of globalizing economic marketplace ideologies and practices (cf. Blommaert 2010).

#### 4. The incursion of English and the rearticulation of space for minoritized languages

In November 2013, the Chilean government, in conjunction with the British Council and the participation of official representatives of fifteen different nations, held a conference titled “Inglés para la prosperidad” (English for prosperity) in Santiago. In the speeches given by officials during the two-day event, two major ideological discursive phenomena were easily discernible: (1) the insertion of English into the social spaces of education and commercial public life – spaces tradi-

tionally occupied by Spanish; and (2) the linkage of English language ability and use with ‘progress’ and participation in the global economy.

Aida Salamanca, the Director of Education and Training Services for the British Council in Colombia, conceptualized metaphorically the role of English, posing the question of its ‘place’ within the Colombian educational realm: “How can we assist the rector in the process of curriculum management so that English is given its due place and is articulated within the school’s plan?”<sup>6</sup> Four fundamental components of the Colombian curriculum (evaluation and monitoring; content and pedagogy; training and mentoring; institutional management support) were illustrated as segments of a large circle in an image projected for the audience, whom Salamanca then asked to imagine how this circle was overlapped by another large circle representing “el sector productivo” (the economic sector). She explained that, because this sector now demands English, the educational sector must respond; the Ministry of Commerce, along with a host of other public and private entities, are fundamentally implicated. Education in Colombia, she insisted, must be conceptualized in its relationship to these other sectors (British Council Chile). Colombian President Juan Manuel Santos’ view is clearly aligned with that expressed by Salamanca; he was recently featured in a national newspaper article teaching a group of local schoolchildren about cities in Colombia in English, in an active show of support for the government’s *Colombia Bilingüe* initiative (*El Espectador* 2015).

Following Salamanca’s explanation, Claudio Seebach, an official of the Ministry of the Secretary General of the Presidency of Chile, described the need for Chile to “adopt” English on a larger scale if it aspires to “progress”, be a “developed” country and maintain its standing in the OECD (Organisation for Economic Co-operation and Development). Interestingly, he clarified that a linguistic command of English is not all that is necessary for the nation to progress but rather the command of a “much larger cultural complex”.<sup>7</sup> He went on to share an anecdote about someone attending the conference whose taxi driver in Santiago on the previous day had been unable to communicate in English. He offered this anecdote as an example of the “global perspective”, which he characterized as “dramatic” in Chile. Seebach’s

---

6. “¿Cómo acompañarle al rector como gestor de los procesos curriculares para que le dé al inglés el lugar que corresponde y lo pueda realmente articular dentro de la programación de la escuela?” (Salamanca, British Council Chile.)

7. “Chile quiere ser un país desarrollado, y esa es la aspiración .... Ser miembro de ese grupo de países [*referring to the OECD*] nos hace aun más preguntarnos, ¿qué son las claves necesarias como país para ser un país desarrollado? Sin duda, el inglés es [una de las claves] ... no solo el manejo del idioma inglés sino el manejo de un complejo cultural mucho más grande ...” (Seebach, British Council Chile).

attempt to naturalize or normalize discursively the due place of English in Latin America – commenting with a tone of explicit dismay that, like the taxi driver from his anecdote, so many people in Chile and other Latin American countries have little or no ability to speak or understand English – is noteworthy. His reference to a “cultural complex” characterized by English language ability is, in my interpretation, related to the sort of capitalist discursive tendencies of which *tuteo* is one aspect (as I explained above). Relying fundamentally upon the metaphor of space and movement, Seebach pointed out that Argentina and the nations of Central America were “going backwards” in terms of English and, concomitantly, in terms of economic development. In Chile, he affirmed, measures were being taken to continue “moving ahead” (British Council Chile).

The notion of English language ability as something ‘natural’ or commonsensical is quickly becoming rooted in the cultural and linguistic imaginaries of Latin America and Spain (cf. Fairclough 1989; Philip, this volume). This attitude/ideology engenders critical subjectivities within language contact settings where the hegemony of Spanish has already been openly debated in public life, educational practice and national language policy. The following quote from a local school-teacher in Peru reveals this tendency:

People are aware that English is the language of science and technology, commerce and international tourism. They tend to incorporate English into their background knowledge. It’s natural. It doesn’t necessarily mean anything against Quechua, but it does make sense that people aspire to learn English.

(Jiménez 2010)<sup>8</sup>

An opinion article published by the Department Head of Humanities, Arts & Social Sciences of the Preparatory School of the Tecnológico de Monterrey in Mexico echoed the same sentiment: “Educational planning for the 21st century must establish the mechanisms for the learning and apprehension of a second language, principally English, regardless of the value of the linguistic legacy that gives cohesion to our indigenous and national cultural heritage and the commitment to preserve it” (Olmedo 2012).<sup>9</sup> The question of whether the acquisition of

---

8. “Las personas están conscientes que el idioma de la ciencia y la tecnología, el comercio y el turismo internacional es el inglés. Tienden a incorporar el inglés a su bagaje. Es natural. No necesariamente tiene que ser en contra del quechua, pero sí tiene sentido que la gente aspire a aprender inglés”.

9. “La planeación del modelo educativo para el siglo XXI, deberá establecer los mecanismos para que una segunda lengua, prioritariamente el inglés, se aprenda y se aprehenda, independientemente del valor y del compromiso por conservar el legado lingüístico que da cohesión al patrimonio cultural indígena y nacional”.



English is positioned *contra* minoritized languages entails a sociocultural phenomenon that is highly characteristic of postmodernity, as described by Hutcheon: “The complicated relationship of the ‘post’ to the ‘modern’ ... is one of critical rethinking, leading either to a continuation and often intensification (of irony, parody, self-reflexivity) or a rejection (of ahistoricity, barriers against the popular)” (2006: 118).

Perhaps paradoxically, the perceived ‘naturalness’ of the social goal to learn English, as expressed by some, appears linked to the ‘natural’ desire to use and maintain the minoritized language of one’s cultural heritage or local identity. Simply stated, recognition of a language other than Spanish compels a sense of self-reflexivity and concomitant agency to maintain local languages, which traditionally have been subordinate to Spanish. This phenomenon manifests itself in recent cultural production in the Peruvian Andes.<sup>10</sup> The pride that young people have increasingly begun to take in indigenous languages throughout the Andean region is evident in the music and visuality of postmodern groups like Cusco’s Proyecto 3399 (whose *Videoarte* series is appropriately titled “Made in Taiwan”). Proyecto 3399’s principal endeavor is remaking in Quechua pop music hits of the past three decades, most of them English-language originals, e.g., Michael Jackson’s “Beat It”; Aqua’s “Barbie Girl”; Britney Spears’ “One More Time”; Lady Gaga’s “Alejandro”. As Jorge Flores, one of the group’s creators, affirmed, “we have a super relationship with things Incan and with our ancestry, but we are also children of Britney Spears, of Lady Gaga, of Madonna, so how do we get people to understand that?” (Tuteve ATV 2013).<sup>11</sup> Flores recalled that, when the group’s remakes were first played in nightclubs in Lima, some people thought the lyrics were being sung in Japanese or Korean, illustrating the unfamiliar or ‘foreign’ effect of hearing Quechua in that type of setting in the Peruvian capital. Interestingly, Cusco native Milagros Palomino, the group’s lead singer and translator of the original lyrics into

---

10. Rotondo (2013) carefully documents the decentralization of audiovisual production in Peru through the creation of alternative spaces and networks for the construction of subjectivities and the commercialization of cultural products beyond the purview of dominant cultural institutions or centers, i.e., Lima in this case. On a related note, Heller (2013) affirms that, “As zones of both authenticity and multilingualism, former peripheries have much to offer in the new circumstances, both in terms of cultural and linguistic goods .... They may also find themselves repositioned with respect to old centres and peripheries more interested in (or required to) participate in global networks and processes than national ones” (22).

11. “[N]osotros tenemos una súper relación con lo inca y con lo ancestral, pero también somos hijos de Britney Spears, de Lady Gaga, de Madonna, entonces ¿cómo hacer entender eso a la gente?”

Quechua, is an English language teacher in Cusco who grew up speaking entirely Spanish and acquired Quechua as a second language. The case of Palomino reflects a dynamic that has begun to unfold in cities and towns throughout the Andes, to wit: the incursion of English in cultural and educational realms places in sharper relief the sociolinguistic situation of local, minoritized languages, with English entering the fray, perhaps paradoxically, as a site of resistance to Spanish language hegemony.<sup>12</sup>

This same dynamic has been easily observable in Catalonia and the Balearic Islands in recent years. In the largest protest in the history of the Balearic Islands, some 90,000 people filled the streets of Palma de Mallorca in late September 2013 to express their opposition to the legislation of mandatory English-language instruction in public schools. The law they were protesting, ushered in by then President José Ramón Bauzá of the Partido Popular, imposed the requirement of a trilingual model of education in which Spanish would occupy one third of the curriculum, Catalan another third and English one third. The new ‘one-third’ slot relegated to English was perceived by many – in particular young adults, university students and those with higher levels of formal education – as an incursion into the arena of Catalan (Aguiló Mora & Lynch 2017). In a television interview (La Sexta Noche), Bauzá justified the introduction of English in the curriculum by postulating that Catalan is actually not the language spoken in the Balearic Islands, but rather Mallorcan, Minorcan, Ibiza and Formenteran.<sup>13</sup> Bauzá appealed to a notion of authenticity in these local variants of Catalan, in the broader sense, by attempting to make Catalan appear as the socially artificial and ideologically intrusive element in the linguistic space of the Balearic Islands. In a discourse strikingly similar to that of Salamanca (in the case of Colombia) and Seebach (in the case of Chile), Bauzá stated that, in response to economic demand, English must be one of three “equal parts” of education in the Balearic Islands. After pointing out that 60% of job opportunities in the Islands currently require English, Bauzá affirmed that he wanted “liberty, not imposition”. Of course, the imposition he referred to was the use of Catalan as the vehicular language of education, which he characterized as “immersion” in Catalan – perhaps an odd

---

12. The names of the two protagonists of the Peruvian-produced series “El cholo Juanito y Richard Douglas”, greatly popular in Bolivia, are likely no coincidence.

13. As the political, economic, and cultural center of the Catalan-speaking world and of the normalization and standardization efforts for the Catalan language, Barcelona, still struggling for autonomy with respect to Madrid, is a prime example of what Pietikäinen & Kelly-Holmes (2013) describe as “peripheries for particular centres and centres for particular peripheries, or on the move from one to the other” (5).

choice of term given the fact that the majority of Balearic Islanders speak a variety of Catalan at home. He emphasized that the third of the curriculum devoted to Catalan in his trilingual model should not reflect the normative, ‘artificial’ variety imported from Barcelona.

What is particularly compelling in the context of both the Catalan- and Quechua-speaking worlds is that, because Spanish-dominant bilinguals represent a majority of the overall population, the normativization and formal acquisition of Catalan and Quechua give rise to institutional and second-language sorts of codes in those languages, perhaps comparable to ‘interlanguages’ in some cases and, in other cases, the ‘H variety’ as defined by Charles Ferguson in his classic theoretical proposal regarding diglossia (1959). As also occurs in the cases of Basque and Guaraní, a highly standardized variety that serves official and classroom purposes appears socially contrived to many Catalan- and Quechua-dominant bilinguals and older monolinguals.<sup>14</sup> Indeed, one of the popular reactions to the remixed music of Proyecto 3399 in Cusco has been that the language sounds artificial. Similarly, the lack of social ‘naturalness’ of Catalan has been a cause for concern in the *Generalitat de Catalunya* over the past decade, particularly with regards to the local ‘native’ Catalan majority’s lack of acceptance and accommodation of the Catalan varieties learned and spoken by immigrants, especially Latin American immigrants and migrants from other Castilian-speaking regions of Spain. As Woolard (2007) has noted, speakers who lack an ‘authentic’ variety of Catalan are often met by a switch to Castilian on the part of native Catalan speakers. This was so much the case that the *Generalitat* launched a promotional campaign in 2005 called “Dóna corda al català” (Wind up Catalan). The aim of the campaign was not only to convince young people to speak more Catalan, but also to convince Catalonians to accommodate the seemingly second-language styles of immigrants and other L2 users of Catalan and to convince L2 users to insist on speaking Catalan despite potential worries about making mistakes or appearing non-native: “Pots parlar en català.

---

14. Lawyer, politician, and member of the Academia de la Lengua Guaraní, Tadeo Zarratea maintains that in order to establish a more stable bilingual competence in Guaraní in Paraguay, the current pedagogical variety must be reconsidered: “Reformar la enseñanza del guaraní. No se puede seguir enseñando este guaraní artificioso de palabras inventadas en gabinete. Enseñar a niños y jóvenes el guaraní hablado, el que está en la garganta del pueblo paraguayano, el que sirve para comunicarse” (Reform the teaching of Guaraní. We cannot keep teaching this artificial Guaraní made up of words invented in an office. Teach children and young people spoken Guaraní, the language which is in the mouths of the Paraguayan people, that which facilitates communication) (Gómez Silguiera 2012).

Parla sense vergonya. Parla amb llibertat” (You can speak Catalan. Speak with no shame. Speak with freedom).<sup>15</sup>

The impetus given to Catalan in recent decades has been simultaneous with the broad social adoption of English as lingua franca in commerce and in the vibrant tourism and service industries of Catalonia and the Balearic Islands. The presence of English has greatly increased in academic settings as well (universities in particular) in the past few years, concomitant with the acute incursion of English as the language of scientific research and publication, technology and Internet communication (cf. Ferguson 2007). Despite the fact that there are more than 500 million Spanish speakers worldwide, in 2013 it was estimated that only 7.8% of Internet users were Spanish speakers (Instituto Cervantes 2013). Pedro García Barreno, member of the *Comisión de Vocabulario Científico y Técnico* of the Real Academia Española (RAE), has attributed this disproportion directly to a lack of economic growth in Latin America, affirming that the “tremendous explosion” of social media has begun to revalorize local languages (Martín Mayorga 2011). Indeed, what some might have considered the geopolitical (i.e., modern) space of Spanish on the Internet has ended up being, in sociolinguistic reality, a much more fragmented (i.e., postmodern) field in which English and other local languages claim territory.

## 5. The normalization of Spanish in cyberspace and global mass media

Perhaps nowhere else in the Spanish-speaking world has the debate regarding language normalization been more intriguing in recent years than in cyberspace and in the Spanish-language mass media emanating from the US. What makes both of these spaces of Spanish language use so compelling is the fact that they transcend geopolitical borders. The hyperliterate populations of postmodern societies access mass media and engage in online communities on an everyday basis. The question of how to conceptualize this linguistic space from the perspective of language variation and change, as well as language contact, poses somewhat of a challenge

---

15. The term *llibertat* has been pervasive in the discourse of language rights in Spain since the democratic transition following Franco’s death. Its connotations are oftentimes not quite those of ‘freedom of speech’ carried by US discursive ideologies, but rather of institutionally sanctioned use and political autonomy of non-Castilian languages, i.e., freedom from the Castilian-normative imposition of the central government in Madrid. Bauzá rather subversively appropriated this trope by implicitly casting Barcelona as the Madrid of the Catalan-speaking world.

to the 20th-century paradigm of sociolinguistics. Internet sites are, by virtue of their existence, everywhere at once, neither really 'here' nor 'there'. They are constituted, represented, constantly and continuously negotiated through imagery and through the written word, i.e., through linguistic form. Indisputably, much of that form on the Internet either looks like English – the so-called global language par excellence – or like English-influenced codes. Perhaps for that reason, most of what has been said about the linguistic dimensions of postmodernity is focused on so-called Global English or on Anglophone contexts. Extremely little has been said about Spanish, also characterized by many as a 'global language' (cf. Del Valle 2007; Mar-Molinero 2000). In a provocative article titled "Spanish, second language of the Internet?" Gabilondo (2006) affirmed that:

For the first time in history, Spanish is a peripheral language [on the Internet], similar to the way in which Basque or Nahuatl became peripheral in their respective states ... [T]he internet brings about a qualitative change in this marginalization [of Spanish], since even translation is no longer an option ... The reference to the periphery ... requires a new imagination that nationalist understandings of culture have made unthinkable for most Spanish users: the Spanish speaker is a peripherally-situated web-user ... The interconnectedness of the web demands the acceptance of the *hybrid* nature of any Spanish use of the web. (123–124)

At the crux of Gabilondo's argument is the dissolution of national, geopolitically defined linguistic borders within which Spanish prevails as the hegemonic language.

As such, incursions of English and English-influenced codes in Spanish-speaking spheres occur on a daily basis for the great majority of Internet users. Informational sites, chat rooms, travel services and online gaming are all prime examples of this phenomenon through which Spanish speakers are exposed to and acquire some use of English, albeit only at the lexical level or in basic discursive repertoires or literacies. Of course, this does not mean that Spanish speakers are becoming English speakers by virtue of their Internet usage. It does mean, however, that, as Gabilondo affirms, within the framework of 20th-century modernity, cyberspace "opens the gates to the linguistic invasion of other languages, mainly English" (2006: 124). What has begun to coalesce in the postmodern context is what I might characterize as more 'English-informed' varieties of Spanish, particularly in Spain and among the urban youth of Latin America. In these emergent varieties, we can readily observe English loanwords – particularly related to popular culture, television and film, fashion, science and technology – and English-influenced lexical and phrasal calques, as well as pragmatic-discursive patterns inspired by Anglophone ways of saying and 'doing things with words'.

This argument is quite similar to one already made by Otheguy & García (1988) regarding variation of Spanish in the US, based on an empirical study of lexical innovations among Cuban immigrants in Miami during the 1980s: “phrasal calques are cases of cultural, not linguistic, innovation” (227). Contrary to expectation, these authors observed that English-based innovations in Miami were more widely diffused among speakers who had more limited knowledge of English, i.e., those who did not know the language of origin of the innovation, than among those who were more bilingual (205); Lynch (forthcoming) documents the persistence of this tendency among Miami Cubans more than three decades later. Otheguy & García concluded that “lexical penetration from English ... is best seen as an instance of creative language adaptation on the part of a community that apparently wishes to continue speaking Spanish while communicating message types that are characteristic of an English-speaking society” (1988: 227). I propose that this argument can be extended on a macro-level scale to the situation of Spanish speakers in postmodern settings throughout Latin America and Spain, especially in cyberspace.

The pervasiveness of English as the code par excellence of global culture and late capitalism, the creation of spaces for previously minoritized local languages on the Internet and the ongoing construction of neoliberal marketplaces are all phenomena that have conspired to place the normativization efforts of the *Academias de la Lengua* in a postmodern bind. As RAE member Pedro García Barreno remarked, “The domain of language was not that of the economy nor that of innovation; the *Academias* were – and in many respects continue to be – structured in such a way that makes it difficult for them to get away from the old way of doing things” (Martín Mayorga 2011).<sup>16</sup> The bind of the *Academia* has become particularly dire in the context of – and in relation to – the US. In what he characterizes as “ideología lingüística mercantil” (an ideology of linguistic commodification), Del Valle (2007) has argued that the efforts of the *Academia* to foment a sense of ‘pan-Hispanic nationalism’ among Spanish speakers in the US are part and parcel of a post-national project of economic imperialism emanating from Spain, principally Madrid (45). According to Del Valle, the ideological linchpin of this late capitalist phenomenon is ‘*hispanofonía*’: a market-based conceptualization of the Spanish-speaking community in which Spanish capital would be perceived as ‘natural’ and ‘legitimate’ (2007: 39).

---

16. “[E]l ámbito de la lengua no era el de la economía ni el de la innovación; las Academias estaban – y en muchos aspectos siguen estando – estructuradas de una manera que dificulta salirse de los viejos guiones”.

In the zeal to naturalize and legitimize Spanish language use in the US, forging a unified '*hispanounidense*' community, the *Academia Norteamericana de la Lengua Española* (ANLE) has found itself in a paradox of the sort so characteristic of postmodernity, to wit: part of ANLE's stated mission on their website is to "ensure that the Spanish used by all US Hispanic speakers will serve to strengthen an AUTHENTIC bilingualism that will enrich the culture of the United States" (emphasis mine).<sup>17</sup> Yet another, seemingly more fundamental, aspect of their stated mission is the following: "To preserve the Spanish language by establishing, diffusing and promoting criteria of PROPRIETY AND CORRECTNESS THROUGH NORMS that justify and clarify its use among US Hispanics" (emphasis is mine).<sup>18</sup> As Lynch & Potowski (2014) suggest, the types of normativizing projects that the ANLE has recently undertaken are rather counterproductive to – and contradictory with – the stated aim of strengthening an "authentic bilingualism" among the US Hispanic population. Several decades of sociolinguistic research on Spanish-English contact in the US have unequivocally demonstrated that English-origin loanwords and English-based calques and pragmatic-discursive patterns are entirely 'natural' features of 'authentic' bilinguals who speak both Spanish and English with high levels of proficiency (Klee & Lynch 2009). In its explicit rejection of these natural language contact phenomena – for the sake of a normative 'US Spanish' – ANLE is unable to reconcile social naturalness and authenticity with linguistic artificiality and anonymity, in the sense explained by Woolard (2007).<sup>19</sup>

The paradoxical dilemma in which ANLE finds itself is rather similar to that faced by US-based Spanish-language television conglomerates such as Univisión and Telemundo in their continuous quest for bigger markets, i.e., greater numbers of viewers and higher rates for advertisers.<sup>20</sup> Spanish-language mass media have seemingly begun to emerge as a sort of center – in postcolonial terms, now

---

17. "procurar que el español usado por todo hispanounidense sirva de base para el fortalecimiento de un bilingüismo auténtico que enriquezca la cultura de los EE.UU".

18. "Preservar el uso de la lengua española estableciendo, difundiendo y promoviendo entre los hispanounidenses los criterios de propiedad y corrección mediante normas que justifiquen y clarifiquen ese uso".

19. Hutcheon notes that postmodernity "as a social and political condition appear[s] fundamentally contradictory, or at the very least paradoxical: it [is] a break from and a continuation of what ha[s] come before" (2006: 121). Some of the key traits of postmodernity highlighted by Hutcheon, in addition to contradiction and paradox, are irony and parody; the displacement of 'metanarratives' by 'micro-' or 'little narratives'; self-reflexivity; being both inward-looking and at the same time outward-looking; both politicized and fence-sitting.

20. See Dávila (2001) on "the making of Latinos" in US mass media spaces, and Dávila & Rivero (2014) for cultural and political perspectives on US Latino/a media.



postmodern terms – for the normativization and normalization of Spanish in the US.<sup>21</sup> The media are thus in the position to construct – or perhaps having to construct – normative codes of Spanish identified as pertinent to, or characteristic of, ‘US Spanish’, a unified sort of geopolitical variety that arguably heretofore has not existed as such (cf. Lynch 2013). Producers, directors, news reporters and actors of US-produced *telenovelas* commonly deem this phenomenon ‘*español neutro*’, sociolinguistically engineered for the purposes of fomenting solidarity within television markets and thus building a larger viewership, both nationally (within the US) and globally (across all of Latin America). Acosta-Alzuru has referred to this as the ‘Miami model’ of US production of Spanish-language *telenovelas* for global export, that is, a show set in a highly visible Hispanic city such as Miami or Los Angeles, with at least one leading role played by a Mexican-origin actor, and a cast of other actors from diverse countries of the Spanish-speaking world who speak with a ‘neutral’ accent (cited in Piñón & Rojas 2011). Most in the industry readily admit that this ‘*español neutro*’ is heavily influenced or based upon normative Mexican Spanish, itself the product of a long tradition of sociolinguistic engineering for purposes of Mexico City’s highly lucrative television and film industry during the 20th century (Valencia & Lynch 2016). Those involved in the US Spanish-language television industry generally opine that it is probably ‘easier’ for Mexicans to adapt to the target variety of ‘*español neutro*’ than it is for those of Caribbean or Argentine origins, for example (see Artman 2015).

Pondering the possibility of a ‘Global Spanish’ recently in *El país*, Morábito affirmed that “not even television, which has always been a potent source of linguistic homogenization, escapes the law of incessant proliferation of localisms, idioms, slang and other nonce and oftentimes ephemeral (though nonetheless significant) uses through which any living language sustains itself”.<sup>22</sup> This phenomenon, which drives programming for niche markets drawn along national immigrant identity lines (e.g., Mexicans in Southern California, Cubans in Miami, Dominicans in New York, etc.), and the fact that US-born Latinos are principally English speakers who prefer programming strictly in English over programming

---

21. Of course, the use of Spanish in the US remains highly variable and discontinuous, particularly among US-born speakers, who overwhelmingly prefer English not only in everyday interactions with their peers but also in the television programming that they choose to watch. In the industry, there is active and ongoing debate about how to best reach English-language Latino television audiences (Piñón & Rojas 2011).

22. “Ni siquiera la televisión, que ha sido siempre un potente factor de homogeneización lingüística, escapa a la ley de la proliferación incesante de localismos, modismos, jergas y demás usos puntuales y a menudo efímeros (y no por efímeros menos significativos) en los cuales se sustenta cualquier lengua viva”.



strictly in Spanish (Piñón & Rojas 2011), are both vestiges of the paradigm of modernity which continue to make US-manufactured ‘pan-Hispanic’ programming in so-called ‘US Spanish’ rather chimerical. Within the mass media, we bear witness to the contradictory postmodern struggle of a “break from and a continuation of what had come before” (Hutcheon 2006: 121), i.e., between the geopolitical delimitations of linguistic space in modernity and the ‘borderless’ transnational reconfiguration of language in postmodernity. In that respect, Univisión, Telemundo and other global cable conglomerates could arguably be for Spanish in the postmodern era of the US what Julióbrega was for Romance in the Iberian Peninsula 2,000 years ago: a mooring for language (cf. Dávila 2001).

## 6. Conclusion

It is noteworthy that, in the arena of mass media and Internet, the market value of Spanish is based on its ubiquity while, in Miami’s global economic workplace, its value seems derived at least in part from its apparent (or feared) scarcity, i.e., US-born generations lack the particular language registers and abilities that the city will need to remain globally competitive (cf. Reksulak et al. 2004). In both of these spaces, however, what is at stake is a globally agreed-upon ‘market-worthy’ variety of the language, for both cultural and commercial purposes. This is precisely the main sociolinguistic issue of concern, and perhaps the most compelling piece of evidence that linguistic frames of modernity have begun to cede to those of postmodernity. In *Modernity at Large*, Appadurai (1996) argues that nations and nationalities as conceived in the modern era, up to the end of the 20th century, are coming rather undone as we enter a late capitalist order of things and as time and space take on different dimensions (cf. Heller & Duchêne 2012). Space – both physical and ideological – is a fundamental dimension of this process and of humanity’s perception, construction and objectification of language.

Offering examples from various geopolitical locations of the present-day Spanish-speaking world, I have highlighted in this chapter a number of ways in which the Spanish language is being spatially reconfigured vis-à-vis English in the global ‘marketplace’ economy and in the expanding logic of late capitalism. I have argued that, throughout Latin America and Spain, the increasing global imperative of English language acquisition and use in the spheres of education, commerce and tourism, and popular culture (particularly cyberspace and mass media) is having the effect of placing traditionally minoritized languages in sharper relief in the linguistic landscape, sometimes generating greater cultural impetus and political advocacy for their use. I have also pointed out that Internet and ‘pan-Hispanic’ global mass media (particularly those emanating from the US) constitute a

novel arena for Spanish-English language contact phenomena and for the socio-linguistic engineering of a ‘neutral’ or ‘global’ variety of Spanish, which is often imagined synonymously as ‘US Spanish’ in US-based media productions. Unlike the spaces of modernity, these postmodern spaces are mostly unbound by normativizing regimes of national character or normalizing tendencies of a geopolitical or ‘local’ community, unfolding in the special circumstances of hyperliteracy and the blurring of lines between written and oral modes (cf. Gee 2012). All of the macro-level phenomena described in this chapter have already begun preparing the ground for micro-level variations in the linguistic and discursive structure of Spanish (e.g., Riccelli, this volume). They bear close watching by those interested in the impact of economic regimes on language use, the societal nature of contact-induced language phenomena and the context-specific processes of linguistic variation and change as ‘glocal’ sorts of solutions emerge from confrontations of the ‘global’ and the ‘local’.

## References

- Academia Norteamericana de la Lengua Española (ANLE). *Nuestra misión*. Retrieved from <http://www.anle.us/87/Nuestra-mision.html>
- Aguiló Mora, Francisca & Andrew Lynch. 2017. “¿Hablas castellano? Do you speak English? o Xerres mallorquí?: Ideologías y actitudes lingüísticas en Mallorca en una era de crisis económica.” *Studies in Hispanic & Lusophone Linguistics* 10.189–223.
- Anderson, Benedict. 1983. *Imagined Communities: Reflections on the Origin and Spread of Nationalism*. London: Verso.
- Appadurai, Arjun. 1996. *Modernity at Large: Cultural Dimensions of Globalization*. Minneapolis: University of Minnesota Press.
- Artman, Hannah. 2015. “The Mexicanization of Walter Cronkite Spanish”. Retrieved from <https://www.youtube.com/watch?v=JjV4z1EsoO4>
- Blommaert, Jan. 2010. *The Sociolinguistics of Globalization*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511845307
- British Council Chile. *English for Prosperity Colombia Chile Panel Plenary IV*. Retrieved from <http://www.youtube.com/watch?v=8n7Ys5WRWBE>
- Brown, Roger & Albert Gilman. 1960. “The Pronouns of Power and Solidarity”. *Style in Language* ed. by Thomas Albert Sebeok, 253–276. Cambridge, Mass.: MIT Press.
- Carter, Phillip & Andrew Lynch. 2015. “Multilingual Miami: Current Trends in Sociolinguistic Research”. *Language and Linguistics Compass* 9.369–385. doi:10.1111/lnc3.12157
- Da Silva, Emanuel, Mireille McLaughlin & Mary Richards. 2007. “Bilingualism and the Globalized New Economy: The Commodification of Language and Identity”. *Bilingualism: A Social Approach* ed. by Monica Heller, 183–206. London: Palgrave Macmillan.
- Dávila, Arlene. 2001. *Latinos Inc.: Marketing and the Making of a People*. Berkeley: University of California Press.
- Dávila, Arlene & Yeidí Rivero, eds. 2014. *Contemporary Latina/o Media: Production, Circulation, Politics*. New York: NYU Press.

- Del Valle, José, ed. 2007. *La lengua, ¿patria común? Ideas e ideologías del español*. Madrid: Vervuert/ Iberoamericana.
- El Espectador*. Redacción Educación. 2015. "Presidente Santos dictó clase de inglés a estudiantes de Armenia" (5 February). Retrieved from <http://www.elespectador.com/noticias/educacion/presidente-santos-dicto-clase-de-ingles-estudiantes-de-articulo-542347>
- Fairclough, Norman. 1989. *Language and Power*. London: Longman.
- Ferguson, Charles. 1959. "Diglossia". *Word* 1.325–340. doi:10.1080/00437956.1959.11659702
- Ferguson, Gibson. 2007. "The Global Spread of English, Scientific Communication and ESP: Questions of Equity, Access and Domain Loss". *Ibérica* 13.7–38.
- Fernández, Mauro. 2003. "Constitución del orden social y desasosiego: Pronombres de segunda persona y fórmulas de tratamiento en español". Paper presented at the Colloquium on Second Person Pronouns and Forms of Address in the European Languages, Instituto Cervantes, Paris, France. Retrieved from [http://cvc.cervantes.es/lengua/coloquio\\_paris/ponencias/fernandez.htm](http://cvc.cervantes.es/lengua/coloquio_paris/ponencias/fernandez.htm)
- Gabilondo, Joseba. 2006. "Spanish, Second Language of the Internet? The Hispanic Web, Subaltern-Hybrid Cultures, and the Neo-liberal Lettered City". *Revista Canadiense de Estudios Hispánicos* 31.107–129.
- Gee, James Paul. 2012. *Social Linguistics and Literacies: Ideology in Discourses*, 4th ed. London: Routledge.
- Gómez Silguiera, Pedro. 2012. "Se debe dejar de enseñar el guaraní artificioso de palabras inventadas". *ABC Color Paraguay* (7 April). Retrieved from <http://www.abc.com.py/edicion-impresa/locales/se-debe-dejar-de-ensenar-el-guarani-artificioso-de-palabras-inventadas-387442.html>
- Hardin, Karol. 2001. *Pragmatics in Persuasive Discourse in Spanish Television Advertising*. Arlington, Texas: SIL International/ University of Texas at Arlington.
- Heller, Monica. 2013. "Repositioning the Multilingual Periphery: Class, Language, and Transnational Markets in Francophone Canada". *Multilingualism and the Periphery* ed. by Sari Pietikäinen & Helen Kelly-Holmes, 17–34. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199945177.003.0002
- Heller, Monica & Alexandre Duchêne. 2012. "Pride and Profit: Changing Discourses of Language, Capital and Nation-State". *Language in Late Capitalism: Pride and Profit* ed. by Alexandre Duchêne & Monica Heller, 1–21. London: Routledge.
- Hutcheon, Linda. 2006. "Postmodernism". *The Routledge Companion to Critical Theory* ed. by Simon Malpas & Paul Wake, 115–126. London: Routledge.
- Instituto Cervantes. 2013. "El español en el mundo. Anuario del Instituto Cervantes". Retrieved from [http://cvc.cervantes.es/lengua/anuario/anuario\\_13/default.htm](http://cvc.cervantes.es/lengua/anuario/anuario_13/default.htm)
- Jiménez, Paul. 2010. "El quechua en el Perú". Retrieved from <http://www.youtube.com/watch?v=DHiqGiOsGtE>
- Klee, Carol & Andrew Lynch. 2009. *El español en contacto con otras lenguas*. Washington: Georgetown University Press.
- La Sexta Noche. 2013. "Bauzá: 'Es una huelga politizada, detrás de la huelga están los sindicatos, el PSOE y los nacionalistas'". *Lasexta.com* (24 November). Retrieved from [http://www.lasexta.com/programas/sexta-noche/noticias/lasexta-noche-programa-deportivo-mas-audiencia-social-sabado\\_2013112400036.html](http://www.lasexta.com/programas/sexta-noche/noticias/lasexta-noche-programa-deportivo-mas-audiencia-social-sabado_2013112400036.html)
- Laguna, Albert. 2010. "Aquí está Álvarez Guedes: Cuban 'Choteo' and the Politics of Play". *Latino Studies* 8.509–531. doi:10.1057/lst.2010.47

- Lakoff, Robin. 1973. "The Logic of Politeness, or, Minding Your P's & Q's." *Papers from the 9th Regional Meeting Chicago Linguistic Society* ed. by Claudia Corum, Thomas Cedric Smith-Stark & Ann Weiser, 292–305. Chicago: Chicago Linguistic Society.
- Lynch, Andrew. 2000. "Spanish-Speaking Miami in Sociolinguistic Perspective: Bilingualism, Recontact, and Language Maintenance among the Cuban-Origin Population". *Research on Spanish in the United States: Linguistic Issues and Challenges* ed. by Ana Roca, 271–283. Somerville, Mass.: Cascadia.
- Lynch, Andrew. 2009. "Expression of Cultural Standing in Miami: Cuban Spanish Discourse about Fidel Castro and Cuba". *Revista Internacional de Lingüística Iberoamericana* 14.21–48.
- Lynch, Andrew. 2013. "Observaciones sobre comunidad y (dis)continuidad en el estudio sociolingüístico del español en Estados Unidos". *El español en los Estados Unidos: E pluribus unum? Enfoques multidisciplinarios* ed. by Domnita Dumitrescu & Gerardo Piña-Rosales, 67–83. New York: Academia Norteamericana de la Lengua Española.
- Lynch, Andrew. Forthcoming. "The Social Diffusion of English-Based Lexical Innovations in Miami Cuban Spanish". *Cuban Spanish Dialectology: Variation, Contact and Change* ed. by Alejandro Cuza. Washington: Georgetown University Press.
- Lynch, Andrew & Kim Potowski. 2014. "La valoración del habla bilingüe en Estados Unidos: Fundamentos sociolingüísticos y pedagógicos en *Hablando bien se entiende la gente*". *Hispania* 97.32–46. doi:10.1353/hpn.2014.0025
- Mar-Molinero, Clare. 2000. *The Politics of Language in the Spanish-Speaking World. From Colonisation to Globalisation*. London: Routledge.
- Martín Mayorga, Daniel. 2011. "La lengua española en la ciencia y la técnica. Una conversación con el Profesor Pedro García Barreno". *Revista Digital Universitaria* 12:8. Retrieved from <http://www.revista.unam.mx/vol.12/num8/art74/#a>
- Morábito, Fabio. 2014. "¿Es posible un español global?" *El país.com* (13 October). Retrieved from [http://cultura.elpais.com/cultura/2014/10/09/babelia/1412856501\\_684428.html](http://cultura.elpais.com/cultura/2014/10/09/babelia/1412856501_684428.html)
- Murillo Medrano, Jorge. 2003. "La cortesía verbal en situaciones de habla en Costa Rica: Hacia la comprensión de la imagen social en su contexto sociocultural". *Actas del primer coloquio del programa EDICE: La perspectiva no etnocentrista de la cortesía: Identidad sociocultural de las comunidades hispanohablantes* ed. by Diana Bravo. Stockholm: University of Stockholm. Retrieved from [www.primercoloquio.edice.org/Actas](http://www.primercoloquio.edice.org/Actas).
- Olmedo, Juan Carlos. 2012. "La enseñanza de idiomas en la dinámica global". *El Universal.mx* (15 May). Retrieved from <http://www.eluniversalmas.com.mx/editoriales/2012/05/58512.php>
- Otheguy, Ricardo & Ofelia García. 1988. "Diffusion of Lexical Innovations in the Spanish of Cuban Americans". *Research Issues and Problems in United States Spanish: Latin American and Southwestern Varieties* ed. by Jacob Ornstein-Galicia, George Green & Dennis Bixler-Márquez, 203–243. Brownsville, Texas: Pan American University.
- Pietikäinen, Sari & Helen Kelly-Holmes. 2013. "Multilingualism and the Periphery". *Multilingualism and the Periphery* ed. by Sari Pietikäinen & Helen Kelly-Holmes, 1–16. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780199945177.003.0001
- Piñón, Juan & Viviana Rojas. 2011. "Language and Cultural Identity in the New Configuration of the US Latino TV Industry". *Global Media and Communication* 7.129–147. doi:10.1177/1742766511410220
- Pujolar, Joan. 2007. "Bilingualism and the Nation-State in the Post-National Era". *Bilingualism: A Social Approach* ed. by Monica Heller, 71–95. London: Palgrave Macmillan.

- Reksulak, Michael, William Shughart & Robert Tollison. 2004. "Economics and English: Language Growth in Economic Perspective". *Southern Economic Journal* 71.232–259. doi:10.2307/4135290
- Rodríguez Solórzano, Adrián. 1999. "Mezcla de tuteo y voseo. El voseo identifica y diferencia al tico". *La Nación Digital* (31 May). Retrieved from [http://www.nacion.com/ln\\_ee/1999/mayo/31/opinion5.html](http://www.nacion.com/ln_ee/1999/mayo/31/opinion5.html)
- Rotondo, Santiago Alfaro. 2013. "Peruwood: La industria del video digital en el Perú". *Latin American Research Review* 48.69–99. doi:10.1353/lar.2013.0049
- Tuteve ATV. 2013. "Conozca '3399', un intento por salvar la extinción del quechua". Retrieved from <http://www.youtube.com/watch?v=BSBzWu4YXJI>
- Valdés, Guadalupe & Cecilia Pino. 1981. "Muy a tus órdenes: Compliment Responses among Mexican-American Bilinguals". *Language in Society* 10.53–72. doi:10.1017/S0047404500008423
- Valencia, Marelys & Andrew Lynch. 2016. "Migraciones mediáticas: la translocación del español en televisoras hispanas de Estados Unidos." *Cuadernos AISPI. Revista de la Associazione Ispanisti Italiani* 8. 171–196.
- Vossler, Karl. 2010 [1932]. *The Spirit of Language in Civilization*. Trans. Oscar Oeser. London: Routledge.
- Woolard, Kathryn. 2007. "La autoridad lingüística y las ideologías de la autenticidad y el anonimato". *La lengua, ¿patria común?* ed. by José Del Valle, 129–142. Madrid: Vervuert/Iberoamericana.

# Female migration and its impact on language choice and use among Afro-Costa Rican women

Lisbeth A. Philip

Loyola University New Orleans

This chapter examines the socio-historical and linguistic impact of two separate historical migratory experiences on a group of 127 Afro-Costa Rican women whose ancestors were, in the vast majority of cases, Black West Indian Immigrants who settled in the province of Limón, Costa Rica. It provides a description of two female migrations, to Costa Rica and later to the United States, to understand their impact on the participants' choice of language within the family and in the affective domains. These domains of language are used as variables for the sociolinguistic analysis of Language maintenance and language shift (LMLS) among the sample in question in Puerto Limón and Siquirres, the two main economic areas of the province of Limón.

## 1. Introduction

Delimiting factors that affect language maintenance or language shift is difficult, because what accounts for speakers' language choice and use in language-contact situations is "determined by a multitude of factors that are difficult to discern and to group" (Stoessel 2002: 94). Both Weinreich (1953) and Fishman (1972) cited Schmidt-Rohr's classification of 'domain of language' to refer to situations or situation types of interactions that occur in particular multilingual settings (Fishman 1972: 80). For instance, if bilinguals choose a given language for a given topic, according to Fishman, that topic pertains to a domain in which that language is dominant for their society or for their sub-group. Fishman (1972) claims that language behavior can be best understood by recognizing the existence of domains that are crucial in the study of LMLS, such as the nine posited by Schmidt-Rohr: the family, the playground and the street, the school, the church, literature, the press, the military, the courts and governmental bureaucracy.

Domains such as the family, religion, the neighborhood and education might have a bearing on whether a language group gives up its mother tongue or keeps it (Weinreich 1953). For instance, the home and family are the last domain to give way to shift (cf. Romero & Sessarego, this volume for further discussion of maintenance of minority vernaculars within the family domain). To support this view, Fishman (1972: 82) claims:

Multilingualism often begins in the family and depends upon it for encouragement if not for protection. In other cases, multilingualism withdraws into the family domain after it has been displaced from other domains in which it was previously encountered.

In some situations, “the aspiration to return to the homeland, however unlikely such a return may be, remains a powerful motivation for first generation immigrants” (Dabène & Moore 1995: 25; Romaine 1995). In other situations, poor L2 skills on the part of first generation immigrant parents, together with their level of education, are contributing factors in the transmission of mother tongue to children, and the same can be said about long lasting relationships with non-family members of the same linguistic and ethnic backgrounds.

Within the family, exogamy is a feature of contact which promotes the rapid development of bilingualism among the new generations. Exogamy “is the most positive indicator of incipient shift and an important mechanism for language shift and assimilation” (Paulston 1987: 35). Paulston (1994: 13) explains that language shift frequently begins with women, manifested in choice of code, of marriage and eventually of the language in which they choose to bring up their children. A possible explanation by Paulston is that women in subordinate positions are more sensitive to issues of power, including the language of power, but she does not find enough evidence to support this.

Domains are often peculiar to a given speech community, and they call for insight into the sociocultural dynamics of particular multilingual settings at particular periods in their history (Fishman 1972: 81). By examining the importance of two female migrations, to Costa Rica and later to the United States, this chapter demonstrates their impact on participants’ choice of language within the family and in the affective domains to determine whether these factors bear a significance in the analysis of LMLS of the heritage language, English, in the Central American country of Costa Rica.

## 2. The initial black West Indian female migration to Costa Rica

The linguistic behavior of the communities of West Indian descent in Costa Rica has been shaped by its demographic and social history, which can be



traced back to the latter decades of the 19th century. In Costa Rica, the necessity of lowering costs by shortening the distances involved in the exportation of coffee prompted the Costa Rican government to consider the construction of ports on the Atlantic coast and railroads to reach these areas (Hernández Cruz 1998). An earlier contract entered into by the government was taken over by an American entrepreneur, Minor Cooper Keith, in 1884. After the arduous completion of the railroad in 1891, Minor Keith moved into the business of banana production, resulting in his founding of the United Fruit Company in 1899.

In the early years of railroad construction and the growth of banana cultivation, the West Indian population came from a variety of places, but the majority were from Jamaica. The initial migration pattern was not based on a traditional structure of kinship and social reproduction in which women play a major role in structuring the immigrants' initial settlements (Putnam 2002). Most men in Limón were far from their mothers, female companions and other female relatives (ibid: 51).

1872 marks the first migration of Jamaican laborers, with the arrival of the first ship from Kingston, Jamaica to the port of Limón.<sup>1</sup> In 1883, census data indicated that approximately 47% of the population in Limón was of Jamaican origin (Casey Gaspar 1979: 235). Little research has dealt with the question of the arrival of women during this period (Hutchison Miller 2002). Putnam (2002: 66) observes that “[u]nprecedented numbers of West Indian women had arrived at the height of the Panama canal-era migration and had stayed to raise families in Limón as the years passed”.

Like their male counterparts, Putnam (2002: 43) adds that the “women came to Limón both directly from the British West Indies and by way of established settlements on the Nicaraguan or Colombian coast”. Some women moved to join family members; others came in response to job opportunities as domestic workers for the White employers of the United Fruit Company (McIlwaine 1997: 42; Hutchison Miller 2002). Single women also began to migrate on their own (Purcell 1982; Putnam 2002). West Indian women played a crucial role in the transmission of customs and values and in the education of their children. Most of them had to generate a source of income by sewing, baking, working as housekeepers and providing services to the community through the church (Municipalidad de Limón 1992).

---

1. *Gaceta Oficial*, December 20, 1872, cited in Duncan & Meléndez, *El negro en Costa Rica*, 2005:86.



## 2.1 Composition of the West Indian household

In Costa Rica, the province of Limón has always had one of the highest rates of ‘free consensual unions’<sup>2</sup> or *uniones libres* at a very early age (Lemistre Pujol 1984: 137). Conjugal patterns in the 1920s census in Costa Rica suggested that marriage and free consensual unions were, as well as households without a male, all part of the lifestyle of the West Indian groups. The 1927 census revealed that childbearing was higher among Black women than White/Hispanic women in Limón in both urban and rural settings.

A common practice among West Indian women was ‘informal fostering’ or ‘child-borrowing’, and they extended this practice to Limón. Conflicting “commitments” and “too many mouths to feed”, as Putnam (2002: 54) puts it, forced some women to entrust the care of their children to either friends or relatives who were in a better financial position.

Among Afro-Costa Rican women, marriage is less frequent in comparison with the general population, and there are many homes which have single mothers as head of the household (Putnam 2004). It is quite common for young women to leave their children with their parents or close relatives in order to work or further their studies in other places. Purcell (1982: 331) indicates that in some cases children stayed in the care of grandparents when their mother established a conjugal union with a man other than their father. Gordon’s (1981: 25) analysis of the West Indian immigrant women applies to this situation:

Caribbean women accept the universal role of childbearer regardless of their socioeconomic and marital status. They also assume the responsibility for the care and protection of their children. Women are often unable to provide the physical care for their children because of the economic roles they must play and therefore must depend upon relatives for such services. The maternal grandmother often serves such a function and she is in turn supported by the daughter.

In the composition of the household where both parents worked, the grandparents also assisted in raising the children and grandchildren. Households with four generations are not uncommon. In fact, one of the participants in this study, PL21, lived with her children in her great-grandparents’ home. Other women in this study either live or have lived with their grandmothers, mothers, aunts, sisters and children under one roof. The presence of the men was minimal to non-existent in some of the homes of the participants interviewed.

---

2. Free consensual unions refer to people living together as husband and wife without going through the formalities of marriage. Such unions did not have legal effect until it was given to them under certain conditions in 1995 (Código de familia de Costa Rica. Título VII. Capítulo Único. Artículos 242–245).

## 2.2 The exclusive enclaves and the gradual increased use of Spanish beginning in the 1930s and 1940s

For an extended period after the initial migration to Costa Rica, West Indians lived in enclaves where English was the language used in the home as well as the workplace (Holm 1983; Viales Hurtado 1998); they were Protestants for the most part, and they wanted to maintain their British educational tradition by building private, often church-sponsored schools. These schools played an important role in the transmission of linguistic and cultural values.

In Limón, speaking English was an advantage, as the main sources of employment – the railroad and banana industry – were owned by American companies whose agents preferred to employ English speakers. Prior to the Second World War, West Indians had little or no need to learn Spanish (Bourgeois 1989, cited in Philip 2012: 219). Spanish was merely an option since it was only required when dealing with government authorities (Harpelle 1993, 2001, cited in Philip 2012: 219).

However, a gradual shift from the exclusive use of English to the incorporation of Spanish took place, when social and educational reforms were implemented to integrate the West Indian Blacks into the national Costa Rican identity (Philip 2012). This marked the passing from monolingualism to bilingualism when the majority of the Black community was mainstreamed into the national identity (Philip 2012). During the 1930s and 1940s the Costa Rican government played a primordial role in the linguistic development of the province. The effect on English language schools was particularly important. By the 1945 academic year, as Philip (2012: 221) explains:

English-speaking students had to abandon their instruction in English upon reaching the 6th grade, due to conflict with their schedules in the Spanish-speaking schools . . . . As a result, maintenance of the heritage language took place during recess time, in the neighborhoods and at home”.

This was the case with SIQ38 and some others of her age group who attended Costa Rican Government schools. According to SIQ40, “Todos hablaban inglés hasta que se fundaron las escuelas de español” (Everyone spoke English until the Spanish schools were founded).<sup>3</sup>

The ‘Spanish’ new curriculum, as Harpelle (2001: 128) indicates, “introduced a history, folklore, and culture that was foreign to Limón and West Indian children, were placed at a further disadvantage”. A new generation of children wanted to learn Spanish to avoid being ridiculed and to facilitate their integration with the

---

3. Unless otherwise noted, all translations are mine.

Spanish-speaking community. However, this brought about serious consequences for them. Speaking Spanish for many became a problem. There was a confusion of sounds and gender in Spanish; many participants still claim that they ‘think’ in English even though their academic training was in Spanish. As explained by PL80 (cited in Philip 2013: 194):

En mi generación, el inglés me ocasionaba problemas en la pronunciación del español, debido a eso les hablo español a los niños, para evitar burlas.

“In my generation, English interfered with my pronunciation in Spanish; for that reason, I speak Spanish to children, to avoid being ridiculed.”

### **3. The effects of the “migration of the wombs” to the United States beginning in the 1940s**

The 1940s and 1950s represented a period of financial hardship for West Indian Blacks in Costa Rica. Many had lost their lands for lack of property titles. Men were not able to sustain their families, and many Afro-Costa Rican women were forced to leave their homes to perform domestic work, outside of the region and abroad. After the right of citizenship was granted, it was seen as the West Indians’ solution to the problem of mobility within Costa Rica, and, as Harpelle (2001: 153) indicates, it gave them protection in their travels abroad. For women, citizenship represented greater choices outside the rural environment, a means of survival and a way of acquiring financial independence and recognition as professionals (Harpelle 1993, 2001). Women’s immigration patterns are determined by economic and social factors. As Friedman Kasaba (2000: 136) explains:

Immigration to developed parts of the world-economy from underdeveloped or peripheral areas has traditionally been argued to improve women’s life chances, and to have generally emancipatory consequences for women in all dimensions of their lives. As women pass through “the golden door” to America, so the story goes, their acquiescence to the demands of traditional patriarchal households and communities (whether a product of coercion or ignorance) is eroded.

Migration of women to the United States in the 1950s and 1960s deserves special attention because it allowed Afro-Costa Rican women to play a major role in the financial survival of their households. During this time some of these women were still undergoing the process of nationality change (Philip 2012). Many women were single mothers, and many of them were in free consensual unions, which facilitated mobility. As a consequence, Limón family structure was significantly altered when many of these women left their families behind in the care of English

speaking grandmothers. The following account of a woman, Laura,<sup>4</sup> depicts an example of the experiences of the Afro-Costa Rican women at this time:

En el caso de mi mamá, mi mamá viajó a los Estados Unidos en esos contratos para hacer oficios allá y mi mamá se fue como a finales del '62 y principios del '63 más o menos. Ella fue con la intención de mejorar su situación económica y de veras, cuando ella fue, pudo mejorar, porque en eso también que estando allá, estando trabajando pues ahorran y enviaban dinero para que nosotros pudiéramos tal vez comprar un lote o mejorar la construcción en la que vivían sus hijos o lo que fuera ... pero el hecho de que emigraran las mujeres para allá era para mejorar su situación económica.

"In the case of my mother, my mother travelled to the United States on those contracts to work there and my mother left more or less by the end of 1962 and beginning of 1963. She left with the intention of improving her financial situation, and truly, when she went, it did improve, because while she was there working, they would save and send money so that we could perhaps purchase land or improve the structure where her children lived or for whatever ... but the fact that women migrated there was to improve their financial situation."

This migration, often referred to as *emigración de los vientres*<sup>5</sup> or "migration of the wombs", contributed to the decline of the Afro-Costa Rican population in the 1960s. Waves of women migrated to places already inhabited by their ethnic peers (mainly New York, New Jersey, Boston and California) where they held a variety of jobs as housekeepers, sitters, nurses and other low-wage jobs.

Not all of the women who migrated to the United States returned to Costa Rica. Those who stayed permanently formed new families while financially supporting the ones left behind in Costa Rica. Many settled legally in the United States, which facilitated the migration of younger Afro-Costa Ricans from the third and fourth generations to join the American labor market, which is more promising than that of Costa Rica (McIlwaine 1997; Duncan 2005). This international migration pattern persists today, based on the narratives of the participants in this study. As Purcell (1982: 244) indicates, "today the families in Limón which do not have close kin or a friend in the United States are few". Afro-Costa Rican Blacks who live in the United States maintain intense correspondence with friends and family (Duncan 2005), perhaps enhanced in today's era in which communication has been facilitated through the Internet. As will be seen in the following section, the history of this community has had a direct effect on its linguistic profile.

4. This is a pseudonym. Data collected via personal communication.

5. This term first appeared in *Luchas y Esperanzas: 100 años de historia doble e inconclusa del cantón de Limón*. Costa Rica: Municipalidad de Limón, 1992, p. 130.

#### 4. Significance of the study

While many claim that age is one of the determining factors associated with language shift in Limón (Spence 1997; Wolfe 2005), special attention should be given to the influence of women in language maintenance and shift, because of their traditional roles as nurturers and language-bearers. “Women make interesting subjects, as women’s speech has been considered conservative, but it is also women who are at the forefront of linguistic change” (Mukherjee 2003: 104). Studies on gender-related language differences are being directed towards women to show their important, if not predominant, role among the multiple factors that cause language maintenance and shift in bilingual communities. One of the most cited works is that of Gal (1979) on the Hungarian speech community in Oberwart, Austria, which documents women’s preference for the use of German over their native Hungarian, based on social prestige and occupational needs (Milroy 1987; Coates 1993; Govindasamy & Nambiar 2003; Mukherjee 2003).

Language choice and use have been examined in various studies on women across the world. In Mexico the process of *castilianization* brought about many changes in the Otomí community to such an extent that, according to Coronado (1984: 96), the Otomí mother is faced with transmitting two languages and the behavioral traits associated with each language to her children: Otomí which she speaks fluently and Spanish which she does not but which is necessary to facilitate the children’s integration in the Spanish speaking environment. Harvey’s study on women in Ocongate, Perú, concluded that Quechua speaking women tend to avoid the active use of Spanish for fear of being socially demoted in their community since they view themselves as “guardians of the indigenous culture and source of moral affection” (1992: 246).

Gonzales-Velázquez’s (1995) article “Sometimes Spanish, sometimes English” examines language choice among three generations of women in the community of Córdova, New Mexico. According to her study, Spanish has been able to survive in this community thanks to their social network of *comadrazgo* and through the role that both family relations and church play in the community.

In Spence’s study (1998: 11, cited in Philip 2013: 200), respondents, who appeared to be female in the majority, displayed both a general admiration for Standard English and a strong support for the maintenance of Limón Creole in the community. However, for many, passing on their language (Limón Creole) to younger generations represents a conflict. Spanish, as the author puts it, “is thought of in terms of status, academic and social advancement”.

## 5. Methodology and sampling

The linguistic situation of the Afro-Costa Rican population is a rather complex one (Philip 2013). At present, the languages spoken by the members of the Afro-Costa Rican population are classified as Spanish; Standard Limón English, which is used in formal settings; and Limón English Creole, also labeled Mekatelyu and Patois, which is used in informal conversations. However, Spanish is the language of literacy, and its coexistence with the English spoken in the area has stimulated speculation about the linguistic vitality of the English spoken in the province of Limón (Philip 2013: 179–180).

In order to obtain an appraisal of the linguistic practices of the group, a sociolinguistic interview by means of a questionnaire was conducted in the two bilingual communities, Puerto Limón and Siquirres, to permit an analysis of some of the variables utilized in the study of LMLS. It is important to add that the methodology and sample employed in this chapter are the same used in a previous publication on attitudes and identity in LMLS by the author (Philip 2013).

Just as in Philip (2013), the sample studied consists of 127 women of West Indian descent who participated in the sociolinguistic interviews. The women were selected based on their availability and willingness to participate in the study. To protect the participants' identities, the labels PLnn and SIQnn were assigned to designate informant interviewed in Puerto Limón and in Siquirres, respectively. The questionnaire was designed in both English and Spanish in order to examine the various factors that account for LMLS within this particular sample. Given the broad array of responses the participants gave for some of the questions, the responses had to be recoded into specific categories to create a smaller number of group scores for the statistical analysis of the data. To illustrate this, the participants' birthplace was recoded into three different groups:

- Born in Costa Rica, outside the province of Limón
- Born in Costa Rica, in the province of Limón
- Born abroad (Jamaica, Nicaragua, Panama, etc.)

Though the questionnaire comprises six parts, the list of sociolinguistic variables tested is as follows: Participants' demographic data (language of interview, location of interview, place of birth, age, profession, marital status); family origins (for the origins of mother, father and maternal and paternal grandparents, the same grouping of the participants' birthplace was applied); participants' language proficiency

and use (including parents' and grandparents' first language/s); language(s) participant speaks, reads and writes;<sup>6</sup> and language use in specific domains.

These variables were tested to determine any association between language choice and use in the analysis of LMLS. The statistical technique used was analysis of variance (ANOVA). The level of significance was set at  $p < .05$ . The responses from the sociolinguistic questionnaire were entered in the software SPSS16.0 for analysis.

## 6. Analysis of data

Of the 127 participants in this study, 71 participants (55.9%) were interviewed in Puerto Limón and 56 (44.1%) were interviewed in the town of Siquirres. In Puerto Limón, 34 participants opted to be interviewed in English whereas 35 chose Spanish; only 2 participants in Limón chose to be interviewed in both languages. In Siquirres, 21 participants chose to be interviewed in English, while 35 participants chose Spanish.

### 6.1 Place of birth of the participants

104 (81.9%) of the 127 participants were born in the province of Limón whereas 15 (11.8%) were born in other provinces of Costa Rica and 8 (6.3%) abroad. Within the province, the city of Puerto Limón was the leading place of birth of the participants in this study; of the 127 participants, 76 of those who were interviewed in both sites reported Puerto Limón as their place of birth.

### 6.2 Age group

In this sample the participants' age group is distributed as follows: 61 are in the 18–45 age group, 45 belong to the 46–65 age group and 21 are in the age group 66 and over.

### 6.3 Education and professional occupations

Of the 125 participants who reported their education level, 17 (13.6%) had completed elementary school while 108 (86.4%) had attended or completed high school, technical/vocational training or university.

---

6. Though it was not a direct question in the questionnaire, in the interviews, I directly asked participants "Did you attend English school?" when they answered *English* (or its varieties) to the question *What language do/did you use at school?* Given the importance of the role of English schools in the area, this was included as a variable for the analysis of LMLS.

The participants in this study come from a wide range of occupations and professions. However, of the 125 participants who reported their occupations, the overwhelming majority, 76 (60.8%), of those in the professional category either work or worked as teachers or in the educational sector. For the remainder, 15 (12.0%) are in the category 'no occupation',<sup>7</sup> 12 (21.6%) work in areas related to personal services and 22 (17.6%) hold jobs in sales and/or office work.

#### 6.4 Marital status

Of the 71 participants interviewed in Puerto Limón, 27 (58.7%) claimed to be single,<sup>8</sup> 28 (56.0%) said that they were married and 16 (51.6%) said they were either separated, divorced or widowed. In Siquirres, of the 56 participants, 19 claimed to be single (41.3%) whereas 22 (44%) said that they were married and 15 (48.4%) said they were either separated, divorced or widowed.

#### 6.5 Family origins of the participants

For the participants of this study, the number of parents born in the province of Limón exceeded those of parents born elsewhere in Costa Rica and abroad. The countries reported for the mothers born abroad were Jamaica, Nicaragua, Panama and the Cayman Islands. On the other hand, the fathers came from a wider variety of places such as St. Lucia, Barbados, Honduras, Belize, Colombia and the United States.

Grandparents born abroad outweighed those born in Costa Rica (both inside and outside of the province). The majority of the participants in this study were of Jamaican ancestry. Other ancestral origins reported were Aruba, Nicaragua, St. Lucia and other French islands, the Cayman Islands, Panama, Barbados, Colombia and Honduras. Germany, England, India and the United States were also noted as the places of birth of some of the participants' grandfathers. The places of origin of the parents and grandparents of the participants of this study is shown in Table 1.

---

7. I included homemakers and students in the category of 'no occupation'. Both active and retired women in domestic services were included under 'personal services'.

8. Some of the women in the 'single' category may have children.



**Table 1.** Origins of participants' parents and grandparents ( $N = 127$ )

Place of birth	<i>N</i>	Frequency	Percentage
<b>Mother</b>	126*		
Costa Rica (outside Limón prov.)		4	3.2%
Costa Rica (in Limón prov.)		89	70.6%
Abroad		33	26.2%
<b>Father</b>	123		
Costa Rica (outside Limón prov.)		10	8.1%
Costa Rica (in Limón prov.)		72	58.5%
Abroad		41	33.3%
<b>Maternal grandmother</b>	119		
Costa Rica (outside Limón prov.)		3	2.5%
Costa Rica (in Limón prov.)		43	36.1%
Abroad		73	61.3%
<b>Maternal grandfather</b>	108		
Costa Rica (outside Limón prov.)		3	2.8%
Costa Rica (in Limón prov.)		30	27.8%
Abroad		75	69.4%
<b>Paternal Grandmother</b>	103		
Costa Rica (outside Limón prov.)		7	6.8%
Costa Rica (in Limón prov.)		35	34.0%
Abroad		61	59.0%
<b>Paternal Grandfather</b>	92		
Costa Rica (outside Limón prov.)		6	6.5%
Costa Rica (in Limón prov.)		25	27.2%
Abroad		61	66.3%

\* ( $N = 127$ ) indicates total number of participants in the study. *N* (second column) represents the actual number of participants for each category presented in the table who responded to the questions (i.e., some questions were not answered by all participants).

## 6.6 Participants' language proficiency

The linguistic proficiency of the people in the area who speak English as L1 is closely linked to the transmission of the language at home and is reinforced by the education received in the English schools. The majority of the participants' grandparents were or are monolingual in English, whereas the majority of their parents tend to be bilingual. As a result, the overwhelming majority of the participants (72.4%) grew up with English as their first language (Table 2).

**Table 2.** Participants' language proficiency ( $N = 127$ )

Participants' language proficiency	<i>N</i>	Frequency	Percentage
<b>Language(s) first learn growing up</b>	127		
Spanish		10	7.9%
Spanish/English		25	19.7%
English		92	72.4%
<b>Language(s) participants speak</b>	127		
Spanish		1	0.8%
Spanish/English		124	98.4%
English		2	1.6%
<b>Language(s) participants read</b>	127		
Spanish		1	0.8%
Spanish/English		124	97.6%
English		2	1.6%
<b>Language(s) participants write</b>			
Spanish	127	16	12.6%
Spanish/English		108	85.0%
English		3	2.4%
<b>Language participants speak most often everyday</b>	126		
Spanish		47	37.3%
Spanish/English		47	37.3%
English		32	25.4%
<b>Attended English schools</b>	95		
No		12	12.6%
Yes		83	87.4%

Of the 95 participants who reported attendance in English schools, only eight in the 18–45 age group reported not receiving any English schooling, while two participants in the 46–65 age group and two in the 66 and above category reported not attending English schools.

## 6.7 Domains of language use: The family

It appears that the family is the most important domain determining the choice of English as the primary language. Even though most of the participants' families are bilingual, as Table 3 indicates, 75 (60%) of the total 125 participants reported English as the language that they speak or spoke with their mothers, and 75 (66.4%) of 113 participants reported that they speak or spoke English with their fathers.

**Table 3.** Domains of language use: the family ( $N = 127$ )

Participants' language use within the family domain	<i>N</i>	Frequency	Percentage
<b>Language(s) used with mother</b>	125		
Spanish		11	8.8%
Spanish/English		39	31.2%
English		75	60.0%
<b>Language(s) used with father</b>	113		
Spanish		12	10.6%
Spanish/English		26	23.0%
English		75	66.4%
<b>Language(s) used with siblings</b>	121		
Spanish		25	20.7%
Spanish/English		51	42.1%
English		45	37.2%
<b>Language(s) used with maternal grandmother</b>	94		
Spanish		6	6.4%
Spanish/English		5	5.3%
English		83	88.3%
<b>Language(s) used with maternal grandfather</b>	67		
Spanish		6	9.0%
Spanish/English		1	1.5%
English		60	89.6%
<b>Language(s) used with paternal grandmother</b>	63		
Spanish		8	12.7%
Spanish/English		3	4.8%
English		52	82.5%
<b>Language(s) used with paternal grandfather</b>	42		
Spanish		7	16.7%
Spanish/English		3	7.1%
English		32	76.2%
<b>Language(s) used with spouse</b>	83		
Spanish		17	20.5%
Spanish/English		34	41.0%
English		32	38.6%

Table 3. (Continued)

Participants' language use within the family domain	N	Frequency	Percentage
<b>Language(s) used with children</b>	99		
Spanish		17	17.2%
Spanish/English		55	55.6%
English		27	27.3%
<b>Language(s) used with immediate relatives</b>	123		
Spanish		18	14.6%
Spanish/English		43	35.0%
English		62	50.4%

In the family domain, the majority of the participants use or used both languages with siblings and spouses. However, of the 99 participants who reported language use with children, 55 (55.6%) reported using both languages with their children. With other members of the family (grandparents and immediate relatives), the majority of the participants reported English as the primary language used.

When comparing participants' responses to the question "Who was the most important family member when you were growing up?" it appeared that female relatives had a larger influence in the upbringing of the participants, with mothers and grandmothers playing a more pivotal role than male relatives, though fathers played a key role for some. Setting aside family structure and dynamics, the family played a crucial role in the linguistic development of the participants in this study. 68.5% of the participants reported that English was the language spoken with the most important family member when growing up. Many participants described their grandparents' knowledge of Spanish as weak and limited to comprehension of spoken Spanish. 26 (20.5%) reported that they used both languages with the most important family member. Only 14 (11%) claimed that they spoke Spanish with the most important family member when growing up.

Now that the participants are adults, for many of them grandparents have become a secondary source of support, being gradually replaced by their parents, especially the mother, who still plays a prominent role. However, for the majority of the participants in both communities, their children have become the most important persons in their families. When the key person is a child, English almost always ceases to be the dominant language spoken with them. Bilingualism becomes the norm among speakers when the most important family member is a child.

6.8 Affective domains

When participants were approached to participate in the interview, many welcomed the author using terms of endearment in Spanish such as in *¿Qué tal mami, en qué la puedo ayudar?* (How's it going, *mami*, how can I help you?). But for some participants, feelings and emotions are best expressed in English. PL54 uses English to express negative things whereas for PL72 English represents the language of literature, poetry and love. Some claim that Spanish is more romantic. In the words of PL56: “Es más rico hablar español cuando está uno enojado” (It is nicer to speak Spanish when one is angry).

In some cases, people tend to codeswitch when they express emotions. The author personally witnessed PL18 complaining in the streets about the garbage situation in the neighborhood: “the [di] damn *basura* man did not *recoger* the [di] damn *basura*”.<sup>9</sup>

Aside from the standard profanity in English, other expressions are considered curse words like “that [dat] bastard”, “that [dat] wretch” (for a woman). In Table 4, of the 73 participants who reported using curse words, 30 (41.1%) stated that they tend to mix both languages when they cursed.

70 participants (61.4%) said that English is used when they do not want anyone to understand what they are saying. One participant stated that she uses a deep Creole whereas another stated that she switches to Jamaican street talk in this type of situation. Another participant stated that when it is time to gossip, she reverts to Mekatelyu, ironically, a form of speech that she cannot tolerate.

Table 4. Affective domain (N = 127)

	N	Frequency	Percentage
Language(s) used when angry	122		
Spanish		17	13.9%
Spanish/English		57	46.7%
English		48	39.3%
Language(s) used when telling jokes	118		
Spanish		48	40.7%
Spanish/English		46	39.0%
English		24	20.3%

9. *Basura* means “garbage”; *recoger* means “to pick up” or “to collect”.

Table 4. (Continued)

	N	Frequency	Percentage
<b>Language(s) used to express affection</b>	127		
Spanish		31	24.4%
Spanish/English		69	54.3%
English		27	21.3%
<b>Language(s) used when praying</b>	126		
Spanish		40	31.7%
Spanish/English		45	35.7%
English		41	32.5%
<b>Language(s) used when cursing</b>	73		
Spanish		14	19.2%
Spanish/English		30	41.1%
English		29	39.7%
<b>Language(s) used when you don't want anyone to understand what you are saying</b>	114		
Spanish		4	3.5%
Spanish/English		40	35.1%
English		70	61.4%

6.9 Demographic factors associated with LMLS

Based on the results of ANOVA, the demographic factors associated with higher maintenance of English were place of interview, language of interview, place of birth and age group. Participants interviewed in Puerto Limón were more likely to maintain the English language than participants interviewed in Siquirres. The language of the interview was highly associated with maintenance of the language. Those who chose to be interviewed in English (L1) were more likely to be maintainers than those who chose Spanish (L2) or both languages. Participants born abroad were more likely to be maintainers. Women 66 and above appeared more likely to be maintainers.

Participants' level of education, marital status and profession/occupation were also considered but were not found to be associated with LMLS. Given the sample size for professions/occupations, the results were not representative factors associated with LMLS.

6.10 Family origins

The origins of parents and both maternal and paternal grandparents were all associated with LMLS. As the mean indicates, participants whose parents and grandparents were born abroad were most likely to be maintainers of L1.

**Table 5.** Participants’ family origin associated with LMLS (*N* = 127)

	<i>N</i>	<i>X</i> (S.D.)	<i>p</i> -value
<b>Mother</b>	126		
Costa Rica (outside Limón prov.)		0.66 (.19)	0.001
Costa Rica (in Limón prov.)		0.71 (.11)	
Abroad		0.80 (.12)	
<b>Father</b>	123		
Costa Rica (outside Limón Prov.)		0.69 (.12)	0.01
Costa Rica (in Limón prov.)		0.71 (.12)	
Abroad		0.78 (.12)	
<b>Maternal grandmother</b>	119		
Costa Rica (outside Limón prov.)		0.50 (.14)	0.001
Costa Rica (in Limón prov.)		0.68 (.12)	
Abroad		0.78 (.10)	
<b>Maternal grandfather</b>	108		
Costa Rica (outside Limón prov.)		0.63 (.22)	0.001
Costa Rica (in Limón prov.)		0.67 (.12)	
Abroad		0.77 (.12)	
<b>Paternal Grandmother</b>	103		
Costa Rica (outside Limón prov.)		0.65 (.12)	0.001
Costa Rica (in Limón prov.)		0.67 (.10)	
Abroad		0.78 (.11)	
<b>Paternal Grandfather</b>	92		
Costa Rica (outside Limón prov.)		0.63 (.13)	0.01
Costa Rica (in Limón prov.)		0.69 (.12)	
Abroad		0.74 (.13)	

*Note.* The mean (*X*) range is 0–1. A mean of 1 indicates most likely to maintain L1.

**6.11    Language proficiency factors**

Parents’ and grandparents’ first language(s) were highly associated with LMLS (Table 8). The effects of language transmission, especially from mothers and maternal grandmothers, together with proficiency in reading and writing learned mainly through the English schools were found to be highly associated with LMLS (*p* = 0.001).

**Table 6.** Participants’ language proficiency associated with LMLS (*N* = 127)

	<i>N</i>	<i>X</i> ( <i>S.D.</i> )	<i>p</i> -value
<b>Language(s) participant speaks</b>	127		
Spanish		0.36*	0.001
Spanish/English		0.73 (.11)	
English		0.98 (.01)	
<b>Language(s) participant reads</b>	127		
Spanish		0.57 (.14)	0.001
Spanish/English		0.74 (.11)	
English		0.95 (0.5)	
<b>Language(s) participant writes</b>	127		
Spanish		0.64 (.14)	0.001
Spanish/English		0.74 (.11)	
English		0.95 (.05)	
<b>Participant’s first language(s)</b>	127		
Spanish		0.53 (.13)	0.001
Spanish/English		0.64 (.08)	
English		0.78 (.09)	
<b>Language(s) used by participant every day</b>	126		
Spanish		0.66 (.12)	0.001
Spanish/English		0.73 (.09)	
English		0.84 (.07)	
<b>Attended English Schools</b>	95		
NO		0.61 (.14)	0.001
YES		0.76 (.10)	

*Note.* The mean (*X*) range is 0–1. A mean of 1 indicates most likely to maintain L1.

\**X* (*S.D.*) = *N* = 1 no *SD* reported.

### 6.12 Factors associated with domains of language use

All factors related to domains of language use were statistically associated with LMLS. Within the family domain, the general tendency for the participants was to use English with immediate family members; however, this was only slightly higher than the propensity towards using both languages, as Table 7 shows.

In the affective domain, the mean scores for English only and English and Spanish speakers were more similar than those who speak Spanish (Table 8). However, participants’ choice of language when they wanted someone to understand what they were saying was significant but not to the extent of the other categories in this domain.



**Table 7.** Domains of language use associated with LMLS: the family domain ( $N = 127$ )

	<i>N</i>	<i>X (S.D.)</i>	<i>p-value</i>
<b>Language(s) used with mother</b>	125		
Spanish		0.50 (.09)	0.001
Spanish/English		0.68 (.08)	
English		0.80 (.08)	
<b>Language(s) used with father</b>	113		
Spanish		0.56 (.13)	0.001
Spanish/English		0.68 (.07)	
English		0.80 (.08)	
<b>Language(s) used with siblings</b>	121		
Spanish		0.60 (.11)	0.001
Spanish/English		0.72 (.07)	
English		0.82 (.08)	
<b>Language(s) used with maternal grandmother</b>	94		
Spanish		0.46 (.07)	0.001
Spanish/English		0.58 (.13)	
English		0.75 (.10)	
<b>Language(s) used with maternal grandfather</b>	67		
Spanish		0.45 (.05)	0.001
Spanish/English		0.66*	
English		0.76 (.10)	
<b>Language(s) used with paternal grandmother</b>	63		
Spanish		0.52 (.10)	0.001
Spanish/English		0.59 (.09)	
English		0.77 (.09)	
<b>Language(s) used with paternal grandfather</b>	42		
Spanish		0.56 (.17)	0.001
Spanish/English		0.60 (.12)	
English		0.78 (.11)	
<b>Language(s) used with spouse</b>	83		
Spanish		0.60 (.12)	0.001
Spanish/English		0.73 (.09)	
English		0.83 (.08)	

Table 7. (Continued)

	<i>N</i>	<i>X</i> (S.D.)	<i>p</i> -value
<b>Language(s) used with children</b>	99		
Spanish		0.59 (.13)	0.001
Spanish/English		0.73 (.08)	
English		0.83 (.09)	
<b>Language(s) used with immediate relatives</b>	123		
Spanish		0.58 (.10)	0.001
Spanish/English		0.72 (.09)	
English		0.79 (.09)	

Note. The mean (*X*) range is 0–1. A mean of 1 indicates most likely to maintain L1.

\* *X* (S.D.) = *N* = 1 no SD reported.

Table 8. Participants' choice and use of languages(s) in the affective domain associated with LMLS (*N* = 127)

	<i>N</i>	<i>X</i> (S.D.)	<i>p</i> -value
<b>Language(s) used when angry</b>	122		
Spanish		0.55 (.12)	0.001
Spanish/English		0.73 (.08)	
English		0.79 (.09)	
<b>Language(s) used to tell jokes</b>	118		
Spanish		0.65 (.10)	0.001
Spanish/English		0.76 (.07)	
English		0.86 (.07)	
<b>Language(s) used to express affection</b>	127		
Spanish		0.64 (.13)	0.001
Spanish/ English		0.74 (.09)	
English		0.82 (.09)	
<b>Language(s) used to pray</b>	126		
Spanish		0.64 (.11)	0.001
Spanish/English		0.74 (.08)	
English		0.82 (.10)	
<b>Language(s) used to curse</b>	73		
Spanish		0.58 (.16)	0.001

(Continued)

Table 8. (Continued)

	N	X (S.D.)	p-value
Spanish/English		0.74 (.10)	
English		0.72 (.13)	
Language(s) used when you don't want others to understand what you are saying	114		
Spanish		0.58 (.21)	0.047
Spanish/English		0.73 (.11)	
English		0.73 (.11)	

Note. The mean (X) range is 0–1. A mean of 1 indicates most likely to maintain L1.

## 7. Discussion

Numerous comments made by the participants indicate that their environment does not lend itself to speaking the heritage language. As Philip (2013: 192) puts it “There seems to be a consensus among the participants that many of the traditions left by the Black West Indians are dying out and consequently this is affecting the fate of their language”. This, despite the fact that the migrations described in this study played a crucial role in the maintenance of their identity, their culture and, of course, the survival of the heritage language. According to SIQ18 (Philip 2013: 192): “Ya se está perdiendo mucho la identidad, los orígenes. Necesitan saber de dónde viene su lengua materna” (We are losing our identity and origins. They need to know where their maternal language comes from). The results obtained from the analysis show that the ability to maintain the English language (L1) is promising in the areas where the interviews were conducted, that is, Puerto Limón and Siquirres, more so in Puerto Limón.

The language of the interview was another factor that demonstrated great influence in LMLS. Participants’ choice of language for the interview is a clear indication of their level of dominance and competency. Even though L1 maintenance was higher among the participants who were interviewed in English in both communities, the participants who opted to be interviewed in Spanish were more numerous than those who were interviewed in English. This is an indication that this sample is characterized by more Spanish-dominant bilinguals.

From the demographic composition of the participants – that is, place of birth, age group, education level, professional occupations and marital status – only place of birth and age group revealed significant information concerning L1 maintenance among the participants. Only those born abroad displayed a higher level of L1 maintenance than the vast majority of participants, who were born in Limón province. The participants born abroad may have come to

Limón province with a high level of proficiency in L1 prior to the interaction with L2, which may corroborate Fishman's (1972) theory on the importance of language maintenance.

For the age variable, the results indicated that L1 maintenance was higher among participants in the 66 and above category. In this study, however, these participants only represented a marginal portion of the sample, not enough to support what Costa Rican researchers (Spence 1997; Wolfe 2005) have expressed about age as one of the determining factors that causes language shift in Limón. A larger sample of older participants would then be required to support Fasold's (1984: 215) statement that "if there is a genuine shift taking place, it would certainly show up in the larger proportions of older speakers using the declining language than younger speakers".

Regarding generational shift theories, Fishman (1972), Dabène & Moore (1995), Vélez (2000) and others have stated that shift undergoes a process of completion by the third and fourth generations. However, most of the participants in this study are third and fourth generation bilinguals, which demonstrates that there is a prolonged stage of bilingualism within this group.

The origins of the participants' parents and grandparents also gave a clear picture of L1 maintenance. It seems that households which are dominated by first generation speakers are more likely to succeed in the maintenance of the language. Maintenance of the L1 is more likely when both parents and grandparents were born abroad than when they were born in Costa Rica. This was observed not only from the results found by establishing their origins but also by the first language spoken. Participants whose parents and grandparents spoke English are more likely to be high L1 maintainers than participants whose parents and grandparents were or are bilingual or monolingual in Spanish.

It seems that the direction of the language toward maintenance or shift becomes more crucial in homes characterized by bilingual families; the effects of shift will intensify in subsequent intergenerational transmission, especially when monolingual (L1) grandparents and parents, who have often acted as L1 reinforcement agents, are absent. Nevertheless, the overwhelming majority of the participants come from bilingual households, some more English-dominant than others, e.g., where participants have stated that it is forbidden to speak Spanish at home.

Participants' language proficiency also revealed that L1 maintenance is higher among the women who were raised with English as their first language. No matter how promising this factor may be, being raised in a certain language does not necessarily guarantee its maintenance. As Fasold (1984: 215) points out, "The young speakers may be monolingual early in life, but it would be foolish to assume that they will remain so all their lives". Nonetheless, maintainers are located among those participants who claim to speak English only, who stated that they read and write in English and use English on a daily basis, although these participants represent a marginal part of this study.

Participants who attended English schools showed more inclination to be L1 maintainers. The English schools, together with the parents and grandparents' first languages, operated as one of the main normative agencies in the maintenance of L1. For most of the participants, reading and writing skills in English came initially through these two channels and perhaps were later reinforced in higher education. As for the influence of the English schools, participants often mentioned that classes in English tended to be more rigorous than those in Spanish.

Although most Afro-Costa Ricans in this study were educated primarily in Spanish, those who received instruction in the English schools still brag that they possess better skills in mathematics than the Spanish-speaking Costa Ricans. Overall, there were many teachers who left behind an educational legacy for the Afro-Costa Rican people in the province of Limón. PL01 was the English teacher of several of the participants, and many others (in both communities) also teach English at the high school level. Participants often claim that if there is a member of the family who teaches or taught English, the pressure to speak only English at home is much greater. For participants PL78 and PL23, this was the case. As PL23 stated during her interview:

Mi familia materna habla inglés. No sólo se habla, sino que se exige. Mi tía es maestra de inglés, ahora reside en los Estados Unidos. La influencia del inglés es más fuerte debido a mi tía.

"My maternal family speaks English. It is not only spoken, it is required. My aunt is an English teacher, now she resides in the United States. The influence of English is stronger because of my aunt."

Today, according to Aguilar-Sánchez (2005), reading skills in English are required for graduation at some major universities in Costa Rica (cf. Lynch, this volume for further discussion of the incursion of English in education in Latin America), and a significant number of these participants hold university degrees. At present, Spanish only and Spanish and English are in competition as the languages spoken by the participants on a daily basis.

The results also demonstrate that the distribution of languages used in social domains in this community support Fishman's (1972) theory that differentiating these domains is key in order to understand the dynamics of language maintenance in any community. As was expected, the family still represents the primary domain where L1 is maintained. High maintenance of L1 was observed with the participants who use or used English within the nucleus of the family, that is, mother, father, siblings, grandparents, spouse, children and immediate relatives. However, for the majority of the participants, the use of L1 is more prevalent with parents, grandparents and immediate family members whereas among siblings, spouses and children the use of both Spanish and English tends to be the norm.

As noted previously, female relatives (grandmothers, aunts, etc.) have provided emotional support and assistance with childcare, and transmission of L1 was also upheld through these channels. Purcell (1982) indicates that maintenance of solidarity is transmitted through close personal and kinship networks. For most women interviewed, L1 came to them through rearing which was reinforced by monolingual (L1) grandparents and parents and by the education received through the English schools.

In the affective domain, the data reveal that L1 maintainers are more likely to be found among the women who use English to express anger or affection, tell jokes and pray. It is interesting to note that cursing is slightly higher among shifters than L1 maintainers. In the affective domain, the trend seems to indicate that English only and the use of both English and Spanish are in competition.

## 8. Conclusion

The functions of the languages in a bilingual community can be analyzed and classified in different ways. Because no set rules that predict the shift in language contact situations have been clearly identified in the scholarly literature, what is important is to determine bilingual speakers' positions with regard to their languages. For Fasold (1984: 239), "language maintenance and shift are the long-term, collective consequences of consistent patterns of language choice". As shown, language choice is the consequence of historical facts that kept the Afro-Costa Rican community marginalized and confronted with racial problems throughout its process of integration (Philip 2012). The legacy left by the two female migrations contributes greatly towards giving a clear picture of the present linguistic situation in terms of language choice and use in the analysis of LMLS.

Duncan (2005) views the family, religion and education as the three main pillars that allowed the Afro-Costa Rican population in Limón to maintain its culture and traditions. This view has also been supported by testimonials published in *Luchas y Esperanzas: 100 años de historia doble e inconclusa del cantón de Limón* (Municipalidad de Limón 1992: 291):

La socialización de los hijos, la transmisión de la cultura, de las tradiciones, de la lengua, corresponde en este contexto social, a las mujeres.

"The socialization of children, the transmission of culture, of traditions, of language in this social context, is the responsibility of women."

Education has been the core value for social mobility for the Afro-Costa Rican population. The upbringing of female offspring comes from a long tradition of the West Indian settlers for whom codes of respectability were very important.

For female children, Purcell (1982) indicates that girls' upbringing was centered on church functions and the English schools' activities. Traditionally, Black West Indian women's work has been associated with domestic services and nurturing roles (McIlwaine 1997). Over time, through the attainment of higher levels of education, perhaps attributed to the second migration, these roles have been replaced with more white-collar professional positions.

Afro-Costa Rican women place great value on education and even greater value on educating their children, especially their daughters. The proportion of women with university degrees is higher among the Afro-Costa Rican than in any other ethnic group except for the Chinese (Putnam 2004). For many members of the Black community, inclusion in the Hispanic spheres, either by speaking Spanish, through education or through intermarriage with Hispanics, was the key to acceptance in 'White' Costa Rican society and, thus, upward social mobility. The influence of the dominant Hispanic language and culture disrupted Afro-Costa Rican culture in such a way as to create cultural and linguistic gaps among the latter.

This study demonstrates that among Afro-Costa Ricans there is a prolonged stage of bilingualism supported by dominant Spanish speakers. Spanish is the language of instruction in schools and the language in which the majority of the Afro-Costa Rican people operate at work, in administrative matters and in their daily interactions with Hispanics. The English spoken in Limón bears no official status in Costa Rica, and, as in many bilingual communities with linguistic minorities, the vitality of the English spoken in the province will depend largely on its peoples' interest and need to preserve it. This interest and need is sometimes greatly stimulated or diminished by economic and social factors and/or attitudes toward the language that will eventually influence the bilingual speaker to choose one language for a given situation as opposed to the other.

## References

- Aguilar-Sánchez, Jorge. 2005. "English in Costa Rica". *World Englishes* 24:2.161–172.  
doi:10.1111/j.1467-971X.2005.00401.x
- Bourgois, Philippe. 1989. *Ethnicity at Work. Divided Labor on a Central American Banana Plantation*. Baltimore, Md.: The Johns Hopkins University Press.
- Casey Gaspar, Jeffrey. 1979. *Limón 1880–1940. Un estudio de la industria bananera en Costa Rica*. San José: Editorial Costa Rica.
- Coates, Jennifer. 1993. *Women, Men and Language: A Sociolinguistic Account of Gender Differences in Language*. London & New York: Longman.
- Coronado Suzán, Gabriela, María Teresa Ramos Enríquez & F. Javier Téllez Ortega. 1984. *Continuidad y cambio en una comunidad bilingüe. Centro de Investigaciones y Estudios Superiores en Antropología Social*. Mexico City: Ediciones de la Casa Chata.

- Dabène, Louise & Danièle Moore. 1995. "Bilingual Speech of Migrant People". *One Speaker, Two languages. Cross-Disciplinary Perspectives on Code-Switching* ed. by Lesley Milroy & Pieter Muysken, 17–44. Cambridge: Cambridge University Press.  
doi:10.1017/CBO9780511620867.002
- Duncan, Quince. 2005. "El negro antillano: inmigración y presencia". *El negro en Costa Rica* ed. by Quince Duncan and Carlos Meléndez, 11th ed., 115–175. San José: Editorial Costa Rica.
- Fasold, Ralph W. 1984. *The Sociolinguistics of Society*. Oxford & New York: Basil Blackwell.
- Fishman, Joshua A. 1972. *Language in Sociocultural Change*. Stanford, Calif.: Stanford University Press.
- Friedman Kasaba, Kathie. 2000. "Ethnic Networks in Women's Migration: A Comparative Study of Jewish and Italian Women in New York, 1870–1924". *Women and Migration. Anthropological Perspectives* ed. by Jacqueline Knörr & Barbara Meier. 135–150. Frankfurt & New York: Campus Verlag & St. Martin's Press.
- Gal, Susan. 1979. *Language Shift. Social Determinants of Linguistic Change in Bilingual Austria*. New York: Academic Press.
- Gonzales Velásquez, María Dolores. 1995. "Sometimes Spanish, Sometimes English. Language Use among Rural New Mexican Chicanas". *Gender Articulated: Language and the Socially Constructed Self* ed. by Kira. Hall & Mary Bucholtz. 421–446. New York: Routledge.
- Gordon, Monica. 1981. "Caribbean Migration: A Perspective on Women". *Female Immigrants to the United States: Caribbean, Latin American, and African Experiences* (RIIES Occasional Paper No. 2., Research Institute of Immigration and Ethnic Studies) ed. by Delores M. Mortimer and Roy S. Bryce-Laporte, 14–55. Washington: Smithsonian Institute.
- Govindasamy, Subra & Mohana Nambiar. 2003. "Social Networks: Applicability to Minority Communities in Multilingual Settings". *International Journal of the Sociology of Language* 161.25–45.
- Harpelle, Ronald N. 1993. "The Social and Political Integration of West Indians in Costa Rica: 1930–50". *Journal of Latin American Studies* 25:1.103–120. doi:10.1017/S0022216X00000389
- Harpelle, Ronald N. 2001. *The West Indians of Costa Rica: Race, Class and Integration of an Ethnic Minority*. Montreal: McGill/Queen's University Press.
- Harvey, Penelope. 1992. "Mujeres que no hablan castellano. Género, poder y bilingüismo en un pueblo andino". *Allpanchis Phuturinga* 38.227–260.
- Hernández Cruz, Omar. 1998. "Culturas y dinámica regional en el Caribe costarricense". *Anuario de Estudios Centroamericanos, Universidad de Costa Rica* 24:1–2.129–162.
- Holm, John, ed. 1983. *Varieties of English Around the World: Central American English*. Heidelberg: Julius Groos Verlag.
- Hutchison Miller, Carmen. 2002. *In Memory of My Ancestors: Contributions of Afro-Jamaican Female Migrants in Port Limón, Costa Rica 1872–1890* (Working Paper Series No. 8). Cave Hill, Barbados: Centre for Gender and Development Studies, University of the West Indies.
- Lemistre Pujol, Annie & Miriam Mayela Acosta Vega. 1984. *Monografía histórica de la provincia de Limón (Costa Rica)*. San José: Editorial Ministerio de Educación Pública.
- McIlwaine, Cathy. 1997. "Vulnerable or Poor? A Study of Ethnic and Gender Disadvantage Among Afro-Caribbeans in Limón, Costa Rica". *The European Journal of Development Research* 9:2.35–61. doi:10.1080/09578819708426689
- Milroy, Lesley. 1987. *Language and Social Networks*. Oxford: Basil Blackwell.
- Mukherjee, Dipika. 2003. "Role of Women in Language Maintenance and Language Shift: Focus on The Bengali Community in Malaysia". *International Journal of the Sociology of Language* 161.103–120.



- Municipalidad de Limón. 1992. *Luchas y Esperanzas: 100 años de historia doble e inconclusa del cantón de Limón*. San José: Uruk Editores, S.A.
- Paulston, Christina Bratt. 1987. "Catalan and Occitan: Comparative Test Cases for a Theory of Language Maintenance and Shift". *International Journal of the Sociology of Language* 63.31–62.
- Paulston, Christina Bratt. 1994. *Linguistic Minorities in Multilingual Settings. Implications for Language Policies*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/sibil.4
- Philip, Lisbeth A. 2012. "From West Indian Black to Afro-Costa Rican: A Socio-Historical Perspective on the Role of Language, Race and Culture in the Transition". *Language and Identity* (Readings in Language Studies, 3) ed. by Paul Chamness Miller, John L. Watzke & Miguel Mantero, 211–231. Grandville, Mich.: International Society for Language Studies, Inc.
- Philip, Lisbeth A. 2013. "Language Maintenance and Language Shift Among Afro-Costa Rican Women". *Language Contact. A Multidimensional Perspective* ed. by Kelechukwu U Ihemere, 177–207. Newcastle upon Tyne: Cambridge Scholars.
- Purcell, Trevor W. 1982. *Conformity and Dissent: Social Inequality, Values and Mobility Among West Indian Migrants in Limon, Costa Rica*. Ph.D. dissertation, John Hopkins University.
- Putnam, Lara. 2002. *The Company They Kept: Migrants and the Politics of Gender in Caribbean Costa Rica, 1870–1960*. Chapel Hill: The University of North Carolina Press.
- Putnam, Lara Elizabeth W. 2004. "La población afrocostarricense según los datos del censo 2000". *Costa Rica a la luz del censo 2000* ed. by Luis Rosero Bixby, 375–398. San José, C.R. Centro Centroamericano de Población de la Universidad de Costa Rica.
- Romaine, Suzanne. 1995. *Bilingualism*. Oxford & Cambridge, Mass.: Blackwell.
- Spence Sharpe, Marva. 1997. "A Case Study of Language Shift in Progress in Port Limon, Costa Rica". *Revista de filología y lingüística de la Universidad de Costa Rica* 23:1.225–234.
- Stoessel, Saskia. 2002. "Investigating the Role of Social Networks in Language Maintenance and Shift". *International Journal of the Sociology of Language* 153.93–131.
- Vélez, Jorge A. 2000. "Understanding Spanish-Language Maintenance in Puerto Rico: Political Will Meets the Demographic Imperative". *International Journal of the Sociology of Language* 142.5–24. doi:10.1515/ijsl.2000.142.5
- Viales Hurtado, Ronny José. 1998. *Después del enclave: un estudio de la Región Atlántica Costarricense*. San José: Editorial de la Universidad de Costa Rica.
- Weinreich, Uriel. 1953. *Languages in Contact. Findings and Problems*. New York: Publications of the Linguistic Circle of New York.
- Wolfe, Terry. 2005. "El porvenir del inglés de Limón". *El negro en Costa Rica* ed. by Quince Duncan and Carlos Meléndez, 11th ed., 246–249. San José: Editorial Costa Rica.

## Hard come, easy go

### Linguistic interfaces in Istanbulite Judeo-Spanish and Afro-Ecuadorian Spanish

Rey Romero<sup>1</sup> & Sandro Sessarego<sup>2,3,4,5</sup>

<sup>1</sup>University of Houston-Downtown / <sup>2</sup>University of Texas at Austin / <sup>3</sup>Freiburg Institute for Advanced Studies / <sup>4</sup>Helsinki Collegium for Advanced Studies /

<sup>5</sup>Foro Latinoamericano de Antropología del Derecho

This study analyzes patterns in syntax-pragmatics (subject pronoun use) and syntax-morphology (nominal and verbal phi-agreement) interfaces in two unrelated varieties of Spanish, namely Afro-Ecuadorian Spanish (AES) and Istanbulite Judeo-Spanish (IJS). In spite of their geographical distance, the traditional variety of older AES speakers and the contemporary variety of younger IJS informants share similar patterns in the aforementioned features. We account for this parallel by proposing a model of cross-generational contact-induced language transmission that is grounded in current theoretical assumptions on the architecture of the faculty of language, its modularity and the vulnerability of modular interfaces. This model suggests that patterns in both traditional AES and contemporary IJS are actually nativized traces of advanced second-language acquisition (SLA) strategies or early first-language attrition (FLA).

#### 1. Introduction

This chapter analyzes a number of grammatical features shared by two dialects of Spanish: Afro-Ecuadorian Spanish (AES) and Istanbulite Judeo-Spanish (IJS) (Romero 2012; Sessarego 2013a). Although these dialects are spoken natively and acquired as L1s in their corresponding populations, these features are also found, to a lesser or greater extent, in L2 and heritage varieties of Spanish: (a) the use of non-emphatic, non-contrastive overt subjects, (b) impoverished nominal, and (c) verbal phi-agreement. The presence of these L2 phenomena in these L1 varieties of Spanish is accounted for through a model of contact-induced language transmission that

builds on recent proposals on the nature of the language faculty (Chomsky 1995), its modular architecture (Jackendoff 1997, 2002; Reinhart 2006; Burkhardt 2005), and its implications for second-language acquisition (SLA) and first-language attrition (FLA) (Rothman & Slabakova 2011; Montrul 2008; Domínguez 2013).

This study focuses on two little-studied contact varieties of Spanish, which have never been in contact with each other, but which nevertheless present similar patterns in the same linguistic domains. This investigation is innovative because it establishes a link between formal approaches to SLA, FLA and the development of certain linguistic phenomena, which put a particular workload on the interfaces between different language modules (i.e., syntax/discourse and syntax/morphology) (White 2009). Section 2 provides a brief historical account of AES and IJS. It describes parallel socio-economic patterns that appear to pressure the two communities under consideration to abandon their traditional variety in favor of the mainstream national language. Section 3 describes the methodology adopted for the collection of the data. Section 4 points to four structural features shared by both AES and IJS. Section 5 consists of a formal analysis of those phenomena. This examination builds on current proposals on the nature of the language faculty and its modularity. Section 6 discusses our findings and proposes a model of cross-generational language transmission capable of accounting for these data. Section 7 concludes the chapter.

## 2. Sociolinguistic similarities between AES and IJS

AES is an Afro-Hispanic variety spoken in northern Ecuador by the descendants of slaves taken to this region to work on Jesuit sugarcane plantations during colonial times (Sessarego 2013a). IJS is spoken by the Sephardic community in Istanbul, Turkey. This Jewish population arrived in Ottoman urban centers after expulsion from Spain in 1492, and it has preserved its variety of Spanish for more than five centuries (Romero 2012).

In spite of their geographical distance, AES and IJS share many commonalities, including sociolinguistic patterns. For instance, both varieties are minority languages, utilized in community-specific linguistic domains for in-group solidarity and cultural identity purposes and rarely employed outside the community. In certain cases, speakers of the mainstream or official languages may present negative attitudes toward the minority varieties, and, since they lack significant use outside the local community, they are perceived as having low value in the linguistic marketplace at the national level (cf. Sankoff & Laberge 1978; cf. Lynch, this volume). Moreover, these minority vernaculars are acquired primarily within the home or family domain, where they are utilized and maintained. In addition, these varieties

may also survive in association with social functions such as religious practices and entertainment such as traditional songs, prayers, etc. (cf. Philip, this volume). For instance, Judeo-Spanish became an integral part of Judaism, and the language was closely linked to Jewish and Sephardic identity, while the most conservative traits of AES are readily found in Chota Valley traditional folkloric songs. Official language and standardization policies in education and professional contexts tend to increase societal pressure to abandon the traditional minority language. Therefore, although minority varieties may play a role in intra-communal economics, ultimately knowledge of the dominant language is needed for economic survival at a larger scale. Communities themselves may encourage children to learn the official or standard variety early on, and children may exhibit failure to fully acquire the minority variety (cf. also Campbell & Muntzel 1989: 189; Romaine 1989: 371).

AES is spoken by the descendants of African slaves who were taken to the Ecuadorian Highlands to work on Jesuit haciendas during the Spanish colonial era (Sessarego 2013a). In 1964, the land reform freed Afro-Ecuadorians from unpaid peonage. This change provided small lots of land for cultivation and access to education to Afro-Ecuadorians, thereby also increasing opportunities for social mobility. The establishment of schools in these rural communities, combined with the possibility of traveling to urban centers for better-paying jobs, resulted in a systematic decrease in the use of AES by Afro-Ecuadorians in favor of the more prestigious regional variety of Spanish, Highland Ecuadorian Spanish (HES) (cf. Sessarego & Gutiérrez-Rexach 2012; Gutiérrez-Rexach & Sessarego 2014). Currently, only a few hundred elderly Afro-Ecuadorians speak a dialect that showcases highly traditional Afro-Hispanic traits, while the younger generations exhibit varieties of Spanish that have almost completely converged with HES (Sessarego 2013a).

After their expulsion from Spain in 1492, the Sephardic Jews settled across the Ottoman Empire, where their Spanish developed into several dialects, many still spoken today (Romero 2012). IJS is the dialect of Spanish spoken by Sephardic Jews in Istanbul, Turkey. Although IJS could have been considered an immigrant language during the first years of settlement in the Ottoman Empire, it was successfully maintained for centuries thanks to the Islamic precept of *zimmi*, which guaranteed Jews autonomy on their religious, educational and legal institutions by paying a poll tax (Sachar 1994: 89–91). Language shift and ultimately language endangerment for IJS occurred after the founding of the Turkish Republic in 1929, when the new government imposed a series of pro-Turkish nationalistic language policies, including compulsory education in Turkish and banning foreign-language schools (Benbassa & Rodrigue 2000: 101–104; Sachar 1994: 101–102). These Turkish-only language campaigns, plus the constant migration of the Sephardic community out of Turkey to Western Europe, the United States, Latin

America and eventually Israel, led to the loss of linguistic domains and a reduced population of speakers in Istanbul. Romero (2012) provides a recent assessment of the current endangered status of IJS, notably as a consequence of these external factors. The younger members of the Istanbulite Jewish community are rapidly shifting to Turkish, showing lower proficiency in IJS than the older generations. Even those members that have managed to maintain IJS are also experiencing phonological and lexical accommodation to Peninsular and Latin American Spanish due to their marketability for international trade and education (Romero 2013).

AES and IJS socio-economic conditions present several similarities. In both linguistic communities, the younger generations tend to shift radically toward the more prestigious language varieties. These common tendencies, nevertheless, lead to opposite results. While the speech of Afro-Ecuadorians in contact with HES gravitates towards more standard varieties of Spanish, since both AES and the regional variety of Spanish are easily accessible to young Afro-Choteños, the Spanish of Istanbulite Sephardic Jews in contact with Turkish exhibits a progressive divergence from the grammatical patterns characteristic of this Romance language. The results of this shifting process lead to a linguistic situation where, to a certain extent, AES and IJS act as ‘mirror images’ of each other. In other words, the patterns of certain features found in the most traditional variety of AES – as spoken by the older generation – precisely match the linguistic phenomena encountered in the speech of the younger IJS speakers. The current study focuses on the parallel linguistic phenomena found in the speech of these two different populations (older AES speakers and younger IJS speakers). We propose a model of contact-induced language transmission that accounts for the shared features of these two Spanish dialects. For ease of presentation, we label these two varieties as traditional AES and contemporary IJS.

### 3. Data collection

The data for AES were collected through sociolinguistic interviews during the winter of 2011–2012. Fifty speakers participated in the fieldwork. All lived in the communities of Tumbabiro, Carpuela, Chota, Santiago, Changuayacu, Chamanal, Concepción, Caldera and Cuajara, nine villages in the Provinces of Imbabura and Carchi, Chota Valley, Ecuador. The informants were native speakers of the local Afro-Hispanic vernacular; they did not speak any other language spoken in the region, such as Quichua. The data for IJS were collected through a series of sociolinguistic interviews in 2007 with more than sixty informants. All of them resided in Istanbul; they could speak the local dialect of Turkish and the variety of IJS they learned at home.

In both cases the interviews were conducted by allowing the speakers to talk about any topic of their liking and by asking them follow-up questions, in line with the principle of Tangential Shift (Labov 1984: 37). The goal was, therefore, to reduce the Observer's Paradox (Labov 1972) as much as possible in order to collect naturalistic data.

#### 4. Shared grammatical features

A comparative analysis of the data collected in these two communities indicated that traditional AES and contemporary IJS share several morphosyntactic features which diverge quite significantly from the grammatical patterns generally found in native varieties of Spanish. These phenomena consist of the overuse of subject pronouns and reduced inflectional morphology across the nominal and verbal domains, as examples (1)–(6) show.

*Use of non-emphatic, non-contrastive overt subjects:*

AES (Sessarego 2013a: 74)

- (1) a. *Nosotros somos de acá porque nosotros vivimo acá desde chico.*  
 we are of here because we lived here since little  
 "We are from here because we lived here since we were kids."
- b. *Yo iba a la ciudad y yo vendía los producto.*  
 I went to the city and I sold the product  
 "I used to go to the city and I used to sell the products."

IJS (Romero 2009: 8)

- (2) a. *Yo digo ke yo la kiero a mi ermuera ke*  
 I say that I her like to my daughter-in-law that  
*está ermoza.*  
 is beautiful  
 "I say that I like my daughter-in-law who is beautiful."
- b. *Tú merkas los gazetos ke tú meldas el día entero.*  
 you buy the newspapers that you read the day all  
 "You buy the newspapers that you read all day long."

*Lack of verbal phi-agreement (person and number):*

AES (Sessarego 2013a: 76)

- (3) a. *Ello dijo que iba al campo.*  
 they said-3SG that went-3SG to the field  
 "They said they were going to the field."
- b. *Cuando yo tuvo uso de razón.*  
 when I had-3SG use of reason  
 "When I was able to think."

IJS (Romero 2009: 28–29)

- (4) a. *Muestras madres dize ke a moz plaze komidas buenas.*  
 our mothers say-3SG that to us like foods good  
 “Our mothers say that we like good food.”
- b. *Los sivdades nuevas es serka de la mar.*  
 the cities new be-3SG near of the sea  
 “The new cities are near the sea.”

*Lack of nominal phi-agreement (number and gender):*

AES (Sessarego 2013a: 70–71)

- (5) a. *Mis hermano joven.*  
 my-PL brother-M.SG young-M.SG  
 “My young brothers.”
- b. *Cuatro hermano joven.*  
 four brother-M.SG young-M.SG  
 “Four young brothers.”
- c. *Todo la cerveza fría.*  
 all-M the-F.SG beer-F.SG cold-F.SG  
 “All the cold beer.”
- d. *Mucho devoción tenían los afro.*  
 much-M.SG devotion-F.SG had the-M.PL afro-M.SG  
 “Africans used to be very devout.”

IJS (Romero 2012: 7, 22)

- (6) a. *Muevos novia están kontente.*  
 new-M.PL bride-F.SG are happy-M.SG  
 “The new brides are happy.”
- b. *Estos ombre son mansevo.*  
 This-M.PL man-M.SG are young-M.SG  
 “These men are young.”
- c. *Los padjinas de los novela de la komunitá.*  
 the-M.PL page-F.PL of the-M.PL news-F.PL of  
 the-F.SG community  
 “I am the editor of the community news pages.”
- d. *Esto es un lingua de sekretos.*  
 This-M.SG is a-M.SG language-F.SG of secrets  
 “This is a language for secrets.”

Interestingly, these phenomena are also found in a number of Afro-Hispanic languages of the Americas (cf. Lipski 2005; Sessarego 2013b), as well as in several heritage dialects of Spanish (Lipski 1985, 1993a; Montrul 2004; etc.), as examples (7–9) show.

*Use of non-emphatic, non-contrastive overt subjects:*

- (7) a. *Yo tando muy pequeña yo conocí a una señora.*  
 I being very small I met to a woman  
 “[When] I was very young I met a woman.”  
 (Afro-Barlovento Spanish, Megenney 1999: 117)
- b. *Claro yo como fue chico yo no acorda vela.*  
 obviously I since was small I no remember candle  
 “Obviously since I was I child I do not remember candles.”  
 (Afro-Bolivian Spanish, Lipski 2008: 101)
- c. *Ella vivía con su mamá y ella quería mucho a su abuelita. Y ella le dijo ...*  
 she lived with her mother and she loved much to her grandmother and she to-her said  
 “She lived with her mother and she loved her granny very much. And she told her ...”  
 (Spanish heritage speaker, Montrul 2004: 133)
- d. *y ella va a visitar a su abuela. Y ella trajo toda la comida para un regalo ...*  
 and she goes to visit to her grandmother and she brought all the food for a present  
 “and she went to visit her grandmother. And she brought all the food as a present ...”  
 (Spanish heritage speaker, Montrul 2004: 134)

*Lack of verbal phi-agreement (person and number):*

- (8) a. *Yo sabe.*  
 I know-3SG  
 “I know.” (Afro-Puertorican, Álvarez Nazario 1974: 194–195)
- b. *Tú habla y no conoce.*  
 you speak-3SG and not know-3SG  
 “You speak and you do not know.” (Afro-Cuban Spanish, Guirao 1938: 3)
- c. *Yo bailo y come.*  
 I dance-1SG and eat-3SG  
 “I dance and I eat.” (Mexican bilinguals in US, Lipski 1993a: 162)
- d. *Viene mis tíos del rancho d’él.*  
 come-3SG my uncles from.the ranch of.him  
 “My uncles are coming from his ranch.”  
 (Mexican bilinguals in US, Lipski 1993a: 162)

*Lack of nominal phi-agreement (number and gender):*

- (9) a. *Tán chiquito puej mij nene.*  
 so little-M.SG then my-PL child-M.SG  
 “So my kids are little.”  
 (Afro-Mexican Oaxacan Spanish, Mayén 2007: 117)



- b. *Gente blanco.*  
people-F.SG white-M.SG  
“White people.” (Cuban Bozal Spanish, Álvarez Nazario 1974: 189)
- c. *Veo a un nariz rojo.*  
I.see to a-M.SG nose-F.SG red-M.SG  
“I see a red nose.” (Spanish heritage speaker, Montrul et al. 2008: 532)
- d. *Tenemos un casa allá.*  
we.have a-M.SG house-F.SG over.there  
“We have a house over there.”  
(Mexican bilinguals in US, Lipski 1993a: 161)

Variation in subject expression in non-emphatic, non-contrastive overt subjects is not uncommon in the Hispanic world, as Table 1 illustrates. The highest percentages of overt subjects tend to be present in the Caribbean geolects and in contact varieties, such as Rivera Spanish in contact with Portuguese (Carvalho & Child 2011) or the several Spanish dialects in contact with English in the United States (e.g., Hochberg 1986).

Table 1. Dialect variation in overall rates of subject expression\*

Geolect	Study	% of overt subjects
Valladolid, Spain	de Prada (2009)	12%
Madrid, Spain	Enríquez (1984)	21%
Mainland immigrants in NYC	Otheguy et al. (2007)	24%
Rivera Spanish (in contact with Portuguese)	Carvalho & Child (2011)	35%
Los Angeles	Silva-Corvalán (1982)	35%
Barranquilla, Colombia	Orozco & Guy (2008)	35%
Puerto Ricans in Boston	Hochberg (1986)	37%
Dominican Spanish in contact with Haitian Creole	Ortiz-López (2010)	57%
San Juan, Puerto Rico	Cameron (1996)	60%
Santo Domingo, D.R.	Cabrera-Puche (2008)	68%
El Cibao, D.R.	Cameron (1996)	70%

\* This table was adapted and modified from Martínez-Sanz (2011: 195).

The higher rates of overt pronouns in the Caribbean area, as well as in the Afro-Hispanic languages of the Americas (AHLAs) (cf. the data on Afro-Peruvian Spanish by Sessarrego & Gutiérrez-Rexach, this volume) have led some researchers to ascribe this phenomenon to a potential creole origin for some of these varieties (cf. Perl & Schwegler 1998). In fact, in several studies, a number of scholars have postulated that the overuse of subject pronouns as well as cases of impoverished

phi-agreement across the DP and CP (among other grammatical phenomena) may be accounted for as the remaining traces of a previous creole language which decreolized subsequent to contact with standard Spanish (cf. de Granda 1970, 1988; Megenney 1993; Álvarez & Obediente 1998; others).

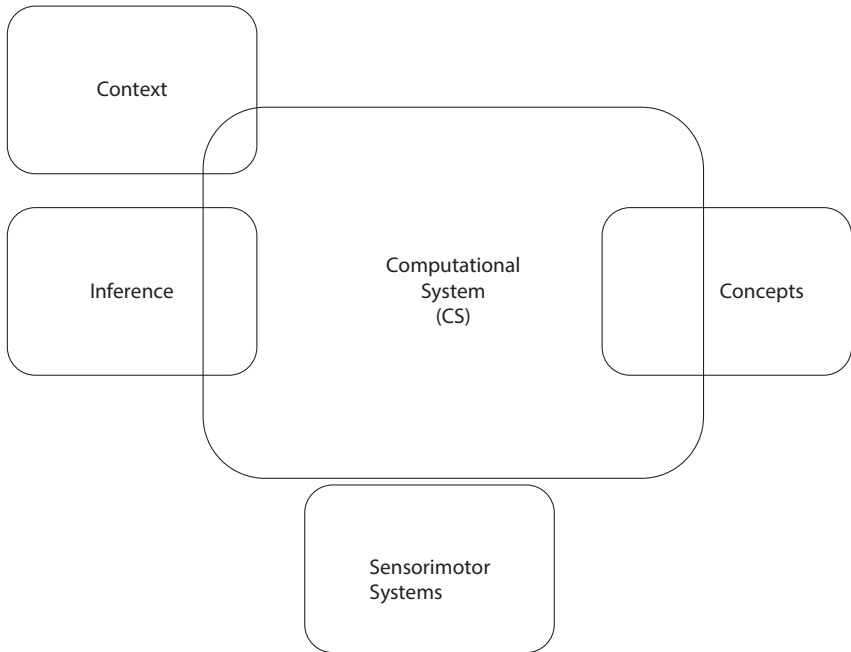
On the other hand, more recent analyses have described the AHLAs as advanced conventionalized interlanguages (Sessarego 2011a, 2011b, 2012, 2013a, 2013b, 2013c, 2014a, 2014b), which should not necessarily be seen as the byproduct of a previous (de)creolization phase (also Lipski 1993b). As for the heritage varieties of Spanish spoken around the world, no one has ever postulated a process of creolization; conversely, attrition and partial language acquisition have often been identified as the processes behind the grammatical nature of such dialects (Montrul 2008). Nevertheless, the presence of identical grammatical patterns in heritage Spanish and in the AHLAs has led some scholars to classify heritage Spanish features as “creoloid” phenomena (Lipski 1985, 1993a). In this chapter, we elaborate on the proposal offered in Sessarego (2013a) and re-elaborated in Sessarego & Gutiérrez-Rexach (this volume) for the diachronic evolution of the AHLAs. We claim that such a model of contact-induced language transmission can be extended to IJS and to other heritage Spanish varieties.

## 5. Data analysis

We propose an analysis of the features reported in examples (1)–(6) that builds on recent theoretical models on the nature of the language faculty, its modularity and the difficulties that certain constructions may pose on the interaction between different language modules. In particular, we adopt here the widely used model of linguistic interface architecture proposed by Reinhart (2006: 3) (Figure 1). It maintains the traditional view on the primacy of syntax, which would interface independently with the other modules forming the language faculty. Although we have adopted Reinhart’s (2006) model, we do not reject alternative proposals on linguistic interfaces (for instance Jackendoff 1997, 2002; Burkhardt 2005, etc.), but rather we leave the decision on the most realistic representation of the faculty of language to further studies.

In recent years several studies on SLA, FLA and bilingualism have focused on the so-called ‘interface hypothesis’ (Sorace 2005; Tsimpli et al. 2004) and its subsequent reformulations (Tsimpli & Sorace 2006; Sorace & Serratrice 2009; Sorace 2011), all of which essentially maintain the core idea that certain constructions involving high processing demand on the interface between different linguistic modules may be more difficult to master in SLA and are the

first to be eroded in cases of FLA. Along these lines, we analyze the features in examples (1)–(6), which hitherto in AHLAs have been ascribed to a previous creole stage and thereby labeled as ‘creolized’. Our approach considers these features as the result of advanced SLA and/or FLA strategies; thus, we suggest that they do not necessarily imply any previous (de)creolization stage for their corresponding dialects.



**Figure 1.** Reinhart's language faculty architecture

The overuse of non-emphatic, non-contrastive overt subjects appears to be linked to the non-target-like acquisition of the null-subject parameter (Chomsky 1981; Rizzi 1982), which characterized the development of AES and IJS. In null-subject languages such as Spanish, subject expression entails the mastery of both syntax and pragmatics and their interface, since both structure and discourse come into play. Typically in these languages, the null subject (*pro*) is employed in topic and non-contrastive focus contexts. Montrul et al. (2009: 303) provide the following example in (10) for the use of *pro* in Spanish:

- (10) *Juan llegó a su casa del trabajo. Primero **pro** se cambió de ropa y luego **pro** decidió ponerse a preparar la cena.*

‘Juan came home from work. First he changed his clothes and then he decided to make dinner.’

Several studies on the L2 acquisition of the null-subject parameter have reported the over-production of overt subjects in contexts where L1 Spanish requires a null subject (White 1985, 1986; Phinney 1987). Additional studies have postulated that L2 learners and early FLA speakers exhibit a surplus of overt subject pronouns because topic features are difficult to acquire, and therefore the native system of overt and covert pronouns is not fully obtained (Sorace 2000, 2003, 2004; Montrul 2004; Domínguez 2013). Indeed, the presence/absence of a [+/- topic shift] feature in a *pro*-drop languages is what accounts for the use of an overt subject (+) or *pro* (-), whereas non-*pro*-drop languages (e.g., English, French) do not have this feature and all subject pronouns must be spelled out. Our data from AES and IJS are in line with the idea that some aspects of the AHLAs and heritage Spanish should be analyzed as SLA/FLA phenomena (Sessarego 2013b). Since the use of *pro* in standard varieties of Spanish entails proficient knowledge of both syntactic and pragmatic features, the overuse of non-emphatic, non-contrastive overt subjects in AES and IJS should not be too surprising (cf. Domínguez 2013). The overuse of overt subjects corresponds to the parallel loss of null subjects in traditional AES and contemporary IJS. This indicates that these varieties underwent a grammatical change which entails a new configuration of the so-called “pro-drop parameter” (Chomsky 1981; Rizzi 1982) (cf. Sessarego & Gutiérrez-Rexach, this volume).

It has been claimed that the linguistic phenomena concerning external interfaces present prolonged challenges for L2 acquisition and L1 attrition, while those involving internal interfaces are more easily mastered (Tsimplici & Sorace 2006; Sorace & Serratrice 2009). However, it has to be said that, even among the realm of internal interfaces, not all of them are created equal. Indeed, some appear to be more problematic than others. The syntax-morphology interface, for example, has been claimed to represent the bottleneck of acquisition (Slabakova 2008, 2009).

In line with Slabakova, we analyze the mastery of phi-features (person, gender, number) as a significant challenge in SLA and FLA, since it puts a high workload demand on the syntax/morphology interface (Slabakova 2008, 2009). Indeed, given that the morphological marking of these features in Spanish is for the most part redundant and does not contribute to the semantic interpretation of the syntactic structure, the complete mastery of such morphemes occurs late in SLA and tends to be easily eroded in FLA (Franceschina 2002; Montrul 2008). In the grammars of Spanish SLA and FLA interlanguages, the incomplete mastery of these features results in morphological variation in agreement across the nominal and verbal domains. Therefore, invariant verb forms for person and number are frequent among SLA and FLA varieties of Spanish, as well as child language (Bybee 1985). In most cases, the use of the third-person

singular appears to be the default form. AES and IJS display variable levels of subject-verb agreement, which in turn reflects their degree of structural change (Sessarego 2012; Romero 2012).

As indicated for Afro-Peruvian Spanish (cf. Sessarego & Gutiérrez-Rexach, this volume), variation in subject-verb agreement for AES and IJS can be formally captured by postulating that these dialects have two potential candidates for Tense Heads (T) to enter the lexical numeration: T1 and T2 (Adger & Smith 2005 present a similar account for Buckie English). In this framework, T1 carries tense, case, number and person features – as in standard varieties of Spanish – while T2 does not bear number and person features. As the operation AGREE (and MERGE) applies between a subject pronoun and T1, the result is a verbal conjugation for tense, number and person. However, when it applies between T2 and its subject, it yields a verbal form conjugated for tense but with default features for number (singular) and person (third person). These operations can be illustrated with the verb *bailar* ‘to dance’ and its subject pronoun *nosotros* ‘we’ in (11) and (12):

- (11) a. T1[tense:PRESENT, ucase:NOM, unum:., upers:] ....  
           pronoun [num:PL, pers:1, ucase:] →
- b. T1[tense:PRESENT, ucase:NOM, unum:PL, upers:1] ....  
           pronoun [num:PL, pers:1, ucase:NOM]
- c. Result: *Nosotros bailamos*  
               we-NOM dance-PRESENT.1PL  
               ‘We are dancing.’
- (12) a. T2 [tense:PRESENT, ucase:NOM] ....  
           pronoun [num:PL, pers:1, ucase:] →
- b. T2 [tense:PRESENT, ucase:NOM] ....  
           pronoun [num:PL, pers:1, ucase:NOM]
- c. Result: *Nosotros baila*  
               we-NOM dance-PRESENT.3SG  
               ‘We are dancing.’

Therefore, the acquisition constraints at the syntax/morphology interface that result in variable subject-verb agreement also help explain the dearth of gender and number features within the nominal domain. Several studies on Spanish SLA and FLA varieties have noted the use of masculine/singular as default values in the DP (White et al. 2004; Sagarra & Herschensohn 2008, 2011; Romero 2011). Moreover, recent research in creolistics postulates a hierarchy in the acquisition of gender/number agreement which initiates in determiners (particularly definite articles) and eventually encompasses other DP elements (cf. Sessarego & Gutiérrez-

Rexach 2011, 2012; Sessarego & Ferreira 2016; Delicado-Cantero & Sessarego 2011; Sessarego 2013c; Sessarego 2014a; Gutiérrez-Rexach & Sessarego 2014). This hierarchy fits well with previous studies in SLA in Romance. For instance, English students acquiring French as a second language exhibited higher agreement percentages in definite than in indefinite articles and, overall, more agreement in determiners than in adjectives. Similar patterns were found for English speakers acquiring Spanish (Bruhn de Garavito & White 2000) and even in Spanish SLA varieties from a wide range of L1 groups such as Italian, Portuguese, Arabic, German and French (Franceschina 2005).

Following the computations previously introduced in (30)–(31) for subject-verb agreement, examples (13)–(14) illustrate concord processes involving gender and number in the DP.

- (13) a. D1[*ugen*;, *unum*:] ... Num[num:PL] ... N1 [gen:F, *unum*:] →  
 b. D1[~~*ugen*~~:F, ~~*unum*~~:PL] ... Num[num:PL] ... N1[gen:F,~~*unum*~~:PL]  
 c. Result: *muchas gatas*  
                   many-F.PL cat-F.PL  
                   “many cats”
- (14) a. D2[ ] ... Num[num:PL] ... N2[gen:F] →  
 b. D2[ ] ... Num[num:PL] ... N2[gen:F]  
 c. Result: *mucho gata*  
                   many-M.SG cat-F.SG  
                   “many cats”

In (13) the determiner (D1) and the noun (N1) come from the lexicon with the standard specification for gender and number features, while in (14) some of those specifications are missing from D2 and N2, thus yielding a different surface result characterized by impoverished agreement (cf. Sessarego 2013c, 2014a for details on impoverished agreement phenomena across the DP).

## 6. The proposal for AES and IJS features

Having presented our analysis, we postulate that the presence of these phenomena in AES and IJS, as well as in virtually all AHLAs and heritage varieties of Spanish, are actually the result of L1 acquisition (‘nativization’) of advanced L2 or early attrition L1 grammars (Sessarego 2013b). This hypothesis follows the premise that L1 and L2 acquisition, as well as L1 attrition, are driven and constrained by Universal Grammar (UG). L1 acquisition is a natural process, an instinct. On the other hand, L2 acquisition and L1 attrition occur differently. L2 development is based on UG, but two inevitable consequences of biological maturation

entail the loss of spontaneity of acquisition and incomplete command of the L2 morpho-lexicon. The interfaces between linguistic modules require high processing for certain features, and L2 speakers may fail to fully acquire them even at advanced stages of L2. These interface constraints also come into play in early L1 attrition, since overuse of overt pronouns and variation of phi-agreement have also been repeatedly attested in FLA varieties (cf. Lipski 1985, 1993a; Romero 2011; Domínguez 2013; others).

Here we adopt a trans-generational scheme of language transmission that can be used as a general model to explain a number of grammatical features encountered in contact varieties (cf. Sessarego 2013b; Sessarego & Gutiérrez-Rexach, this volume). It offers an account for the parallel constructions detected in AES and IJS. Indeed, the presence of identical features in AES, IJS and other AHLAs and heritage Spanish dialects leads us to postulate that, at a certain point in the development of these vernaculars, their speakers (Generation 1) could achieve a level of proficiency in Spanish that closely approximated the TL. Such a phenomenon implied that each speaker was able to acquire a grammar (G), out of a set of possible grammars (G1, G2, Gn) which diverged – to different extents – from the TL. Their linguistic outputs (x, y, z) represented the primary linguistic data (PLD) for the following generation (Generation 2), which acquired such a variety as an L1.

Example (15) depicts the aforementioned scheme where it is possible to observe how Grammar 1 and Grammar 2 (G1 and G2) consist of two linguistic systems with dissimilar parametric configurations:

- (15) a. Individual from Generation 1: TLy → UG driving L2  
acquisition/ L1 attrition → G1 → set of outputs X
- b. Individual from Generation 2: PLDx → UG driving L1  
acquisition → G2 → set of outputs Z

In (15) it is possible to evince how the L1 acquisition processes that lead to G2 are based on PLDx, which are rooted in Generation 1's processes of L2 acquisition/ L1 attrition. Such a trans-generational linguistic transmission is what we classify as 'nativization'. Given this analysis, we provide an explanation for the presence of advanced L2/attrition L1 features in an L1 grammar (G2). This allows us to understand why traditional AES and contemporary IJS exhibit the exact same deviations from Spanish, which are rooted in L2 and attrition phenomena. This type of trans-generational language transmission may also be seen as what has recently been labeled as "linguistic intra-community recycling" (Lipski 2016), where non-target like features are passed on from one generation to the other, thus resulting in a cyclical and systematic shift away from the original native variety.

## 7. Conclusion

This chapter has focused on the nature of some grammatical features shared by two geographically distant Spanish dialects, AES and IJS. We explained such a linguistic parallelism by proposing a model of cross-generational contact-induced language transmission that relies on current assumptions on the architecture of the faculty of language, its modularity and the vulnerability of constructions exerting costly processing workload on its modular interfaces. With this premise, we claim that the linguistic features in question found in traditional AES and contemporary IJS may be analyzed as nativized traces of advanced SLA or early FLA strategies.

## References

- Adger, David & Jennifer Smith. 2005. "Variation and the Minimalist Program". *Syntax and Variation. Reconciling the Biological and the Social* ed. by Leonie Cornips & Karen P. Corrigan, 149–178. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/cilt.265.1oadg
- Álvarez, Alexandra & Enrique Obediente. 1998. "El español caribeño: Antecedentes socio-históricos y lingüísticos". *América negra: Panorámica actual de los estudios lingüísticos sobre variedades hispanas, portuguesas y criollas* ed. by Matthias Perl & Armin Schwegler, 40–61. Madrid & Frankfurt: Iberoamericana/Vervuert.
- Álvarez Nazario, Manuel. 1974. *El elemento afronegroide en el español de Puerto Rico*, 2nd ed. San Juan: Instituto de Cultura Puertorriqueña.
- Benbassa, Esther & Aron Rodrigue. 2000. *Sephardi Jewry: A History of the Judeo-Spanish Community, 14th–20th Centuries*. Los Angeles: University of California Press.
- Bruhn de Garavito, Joyce & Lydia White. 2000. "L2 Acquisition of Spanish DPs: The Status of Grammatical Features". *Proceedings of the 24th Annual Boston University Conference on Language Development* ed. by Catherine Howell, Sarah A. Fish & Thea Keith Lucas, 164–175. Somerville, Mass.: Cascadilla.
- Burkhardt, Petra. 2005. *The Syntax-Discourse Interface: Representing and Interpreting Dependency*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/la.80
- Bybee, Joan L. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/tsl.9
- Cabrera-Puche, María J. 2008. *Null Subject Patterns in Language Contact: The Case of Dominican Spanish*. Ph.D. dissertation, University of Michigan.
- Cameron, Richard. 1996. "A Community-Based Test of a Linguistic Hypothesis". *Language in Society* 25.61–111. doi:10.1017/S0047404500020431
- Campbell, Lyle & Martha C. Muntzel. 1989. "The Structural Consequences of Language Death". *Investigating Obsolescence: Studies in Language Contraction and Death* ed. by Nancy C. Dorian, 181–196. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511620997.016
- Carvalho, Ana M. & Michael Child. 2011. "Subject Pronoun Expression in a Variety of Spanish in Contact with Portuguese". *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics* ed. by Jim Michnowicz & Robin Dodsworth, 14–25. Somerville, Mass.: Cascadilla Proceedings Project.



- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- de Granda, Germán. 1970. "Un temprano testimonio sobre las hablas 'criollas' en África y América". *Thesaurus* 25:1.1–11.
- de Prada, Ana. 2009. *Subject Expression in Minorcan Spanish: Consequences of Contact with Catalan*. Ph.D. dissertation, The Pennsylvania State University.
- Delicado-Cantero, Manuel & Sandro Sessarego. 2011. "Variation and Syntax in Number Expression in Afro-Bolivian Spanish". *Proceedings of the 13th Hispanic Linguistic Symposium* ed. by Luis A. Ortiz-López, 42–53. Somerville, Mass.: Cascadilla.
- Domínguez, Laura. 2013. *Understanding Interfaces*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/lald.55
- Enríquez, Emilia. 1984. *El pronombre personal sujeto en la lengua española hablada en Madrid*. Madrid: Consejo Superior de Investigaciones Científicas.
- Franceschina, Florencia. 2002. "Case and  $\phi$ -Feature Agreement in Advanced L2 Spanish Grammars". *EUROSLA Yearbook* ed. by Susan H. Foster-Cohen, Tanja Ruthenberg & Marie Louise Poschen, 71–86. Herndon, Va.: John Benjamins.
- Franceschina, Florencia. 2005. *Fossilized Second Language Grammars*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/lald.38
- Guirao, Ramón. 1938. *Orbita de la poesía afrocubana, 1928–1937*. Havana: Ucar García.
- Gutiérrez-Rexach, Javier & Sandro Sessarego. 2014. "Morphosyntactic Variation in Three Afro-Andean Dialects: The Evolution of Gender Agreement in DP". *Lingua*. 142–161. doi:10.1016/j.lingua.2014.04.011
- Hochberg, Judith. 1986. "Functional Compensation for /-s/ Deletion in Puerto Rican Spanish". *Language* 62.609–621. doi:10.1353/lan.1986.0041
- Jackendoff, Ray. 1997. *The Architecture of the Language Faculty*. Cambridge, Mass.: MIT Press.
- Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press. doi:10.1093/acprof:oso/9780198270126.001.0001
- Labov, William. 1972. *Language in the Inner City*. Philadelphia: University of Pennsylvania.
- Labov, William. 1984. "Field Methods of the Project on Linguistic Change and Variation". *Language in Use* ed. by John Baugh & Joel Sherzer, 84–112. Englewood Cliffs: Prentice Hall.
- Lipski, John. 1985. "Creole Spanish and Vestigial Spanish: Evolutionary Parallels". *Linguistics* 23.963–984. doi:10.1515/ling.1985.23.6.963
- Lipski, John. 1993a. "Creoloid Phenomena in the Spanish of Transitional Bilinguals". *Spanish in the United States: Linguistic Contact and Diversity* ed. by John Lipski & Ana Roca, 155–182. Berlin: Mouton de Gruyter. doi:10.1515/9783110804973.155
- Lipski, John. 1993b. "On the Non-Creole Bases of Afro-Caribbean Spanish". *Research Paper Series No. 24*. Albuquerque, New Mex.: University of New Mexico.
- Lipski, John. 2005. *A History of Afro-Hispanic Language: Five Centuries, Five Continents*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511627811
- Lipski, John. 2008. *Afro-Bolivian Spanish*. Madrid & Frankfurt: Iberoamericana/Vervuert.
- Lipski, John. 2016. "On the Tenacity of Andean Spanish: Intra-Community Recycling". *Spanish Language and Sociolinguistic Analysis* ed. by Sandro Sessarego & Fernando Tejedo 109–134. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/ihl.8.04lip
- Martínez-Sanz, Cristina. 2011. *Null and Overt Subjects in a Variable System: The Case of Dominican Spanish*. Ph.D. dissertation, University of Ottawa.
- Mayén, Norma. 2007. *Afro-Hispanic Linguistic Remnants in Mexico: The Case of the Costa Chica Region of Oaxaca*. Ph.D. dissertation, Purdue University.

- Megenney, William. 1993. "Elementos criollo-portugueses en el español dominicano". *Montalbán* 15.3–56.
- Megenney, William. 1999. *Aspectos del lenguaje afronegroide en Venezuela*. Madrid & Frankfurt: Iberoamericana/Vervuert.
- Montrul, Silvina. 2004. "Subject and Object Expression in Spanish Heritage Speakers: A Case of Morphosyntactic Convergence". *Bilingualism: Language and Cognition* 7:2.125–142. doi:10.1017/S1366728904001464
- Montrul, Silvina. 2008. *Incomplete Acquisition in Bilingualism. Re-examining the Age Factor*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/sibil.39
- Montrul, Silvina, Rebecca Foote & Silvia Perpiñán. 2008. "Gender Agreement in Adult Second Language Learners and Spanish Heritage Speakers: The Effects of Age and Context of Acquisition". *Language Learning* 58:3.503–553. doi:10.1111/j.1467-9922.2008.00449.x
- Montrul, Silvina, Rejane Dias & Ana Thomé-Williams. 2009. "Subject Expression in the Non-Native Acquisition of Brazilian Portuguese". *Minimalist Inquiries into Child and Adult Language Acquisition: Case Studies across Portuguese* ed. by Acrisio Pires & Jason Rothman, 301–325. Berlin: Mouton DeGruyter. doi:10.1515/9783110215359.2.301
- Orozco, Rafael & Gregory Guy. 2008. "El uso variable de los pronombres sujetos: ¿Qué pasa en la costa caribe colombiana?" *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics* ed. by Maurice Westmoreland & Juan Antonio Thomas, 70–80. Somerville, Mass.: Cascadilla.
- Ortiz-López, Luis. 2010. *El español y el criollo dominicano*. Madrid & Frankfurt: Iberoamericana/Vervuert.
- Otheguy, Ricardo, Ana Celia Zentella & David Livert. 2007. "Language and Dialect Contact in Spanish in New York: Towards the Formation of a Speech Community". *Language* 83:4.770–802. doi:10.1353/lan.2008.0019
- Perl, Matthias & Armin Schwegler, eds. 1998. *América negra: Panorámica actual de los estudios lingüísticos sobre variedades hispanas, portuguesas y criollas*. Madrid & Frankfurt: Iberoamericana/Vervuert.
- Phinney, Marianne. 1987. "The Pro-Drop Parameter in Second Language Acquisition". *Parameter Setting* ed. by Thomas Roeper & Edwin Williams, 221–238. Dordrecht: Reidel. doi:10.1007/978-94-009-3727-7\_10
- Reinhart, Tanya. 2006. *Interface Strategies: Reference-Set Computation*. Cambridge, Mass.: MIT Press.
- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris. doi:10.1515/9783110883718
- Romaine, Suzanne. 1989. "Pidgins, Creoles, Immigrant and Dying Languages". *Investigating Obsolescence: Studies in Language Contraction and Death* ed. by Nancy C. Dorian, 369–383. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511620997.029
- Romero, Rey. 2009. *Corpus of Istanbulite Judeo-Spanish speech*. Unpublished raw data.
- Romero, Rey. 2011. "Mainland vs. Island: A Comparative Morphological Study on Spanish-Turkish Contact". *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics* ed. by Jim Michnowicz & Robin Dodsworth, 50–56. Somerville, Mass.: Cascadilla Press.
- Romero, Rey. 2012. *Spanish in the Bosphorus: A Sociolinguistic Study on the Judeo-Spanish Dialect Spoken in Istanbul*. Istanbul: Libra.
- Romero, Rey. 2013. "Palatal East Meets Velar West: Dialect Contact and Phonological Accommodation in Judeo-Spanish". *Studies in Hispanic and Lusophone Linguistics* 6:2.279–299. doi:10.1515/shll-2013-1151
- Rothman, Jason & Roumyana Slabakova. 2011. "The Mind-Context Divide: On Acquisition at the Linguistic interfaces". *Lingua* 121.568–576. doi:10.1016/j.lingua.2011.01.003

- Sachar, Howard M. 1994. *Farewell España: The World of the Sephardim Remembered*. New York: Vintage.
- Sagarra, Nura & Julia Herschensohn. 2008. "Processing Gender in L2 Spanish". *Proceedings of the 32nd Annual Boston University Conference on Language Development* ed. by Harvey Chan, Heather Jacob & Enkeleida Kapia, 427–437. Somerville, Mass.: Cascadia.
- Sagarra, Nura & Julia Herschensohn. 2011. "Asymmetries in Gender and Number Agreement Processing in Late Bilinguals". *Selected Proceedings of the 13th Hispanic Linguistics Symposium* ed. by Luis A. Ortiz-López, 169–177. Somerville, Mass.: Cascadia.
- Sankoff, David & Suzanne Laberge. 1978. "The Linguistic Market and the Statistical Explanation of Variability". *Linguistic Variation: Models and Methods* ed. by David Sankoff, 239–250. New York: Academic Press.
- Sessarego, Sandro. 2011a. *Introducción al idioma afroboliviano: Una conversación con el Awicho Manuel Barra*. Cochabamba & La Paz: Plural Editores.
- Sessarego, Sandro. 2011b. "On the Status of Afro-Bolivian Spanish Features: Decreolization or Vernacular Universals?" *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics* ed. by Jim Michnowicz & Robin Dodsworth, 125–141. Somerville, Mass.: Cascadia.
- Sessarego, Sandro. 2012. "Non-Creole Features in the Verb System of Afro-Hispanic Languages: New Insights from SLA Studies". *International Journal of Linguistics* 4:1.146–157. doi:10.5296/ijl.v4i1.1415
- Sessarego, Sandro. 2013a. *Chota Valley Spanish*. Madrid & Frankfurt: Iberoamericana/Vervuert.
- Sessarego, Sandro. 2013b. "Afro-Hispanic Contact Varieties as Conventionalized Advanced Second Languages". *IBERIA* 5:1.96–122.
- Sessarego, Sandro. 2013c. "Enhancing Dialogue between Quantitative Sociolinguistics and Minimalist Syntax". *Studies in Hispanic and Lusophone Linguistics* 5:2.79–97.
- Sessarego, Sandro. 2014a. *The Afro-Bolivian Spanish Determiner Phrase: A Microparametric Account*. Columbus, Ohio: The Ohio State University Press.
- Sessarego, Sandro. 2014b. "Afro-Peruvian Spanish in the Context of Spanish Creole Genesis". *Spanish in Context* 11:3.381–401. doi:10.1075/sic.11.3.04ses
- Sessarego, Sandro & Javier Gutiérrez-Rexach. 2011. "A Minimalist Approach to Gender Agreement in the Afro-Bolivian DP: Variation and the Specification of Uninterpretable Features". *The Diachronic of Gender Marking* (Special Issue of *Folia Linguistica*) 45:2 ed. by Gunther de Vogelaer & Mark Janse, 465–488.
- Sessarego, Sandro & Javier Gutiérrez-Rexach. 2012. "Variation, Universals, and Contact-Induced Change: Language Evolution across Generations and Domains". *Current Formal Aspects of Spanish Syntax and Semantics* ed. by Melvin González-Rivera, 251–270. Newcastle: Cambridge Scholars.
- Sessarego, Sandro & Letania Ferreira. 2016. "Spanish and Portuguese Parallels: Impoverished Number Agreement as a Vernacular Feature of Two Rural Dialects". *Spanish Language and Sociolinguistic Analysis* ed. by Sandro Sessarego & Fernando Tejado, 283–304. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/ihll.8.12ses
- Silva-Corvalán, Carmen. 1982. "Subject Expression and Placement in Mexican-American Spanish". *Spanish in the United States: Sociolinguistic Aspects* ed. by Jon Amastae & Lucía Elías-Olivares, 93–120. Cambridge & New York: Cambridge University Press.
- Slabakova, Roumyana. 2008. *Meaning in the Second Language*. Berlin: Mouton de Gruyter. doi:10.1515/9783110211511
- Slabakova, Roumyana. 2009. "What is Easy and What is Hard to Acquire in a Second Language?" *Proceedings of the 10th Generative Approaches to Second Language Acquisition*

- Conference (GASLA 2009)* ed. by Melissa Bowles, Tania Ionin, Silvina Montrul & Annie Tremblay, 280–294. Somerville, Mass.: Cascadilla.
- Sorace, Antonella. 2000. “Syntactic Optionality in Non-Native Grammars”. *Second Language Research* 16.93–102. doi:10.1191/026765800670666032
- Sorace, Antonella. 2003. “Near-Nativeness”. *The Handbook of Second Language Acquisition* ed. by Catherine J. Doughty & Michael H. Long, 130–153. Oxford: Blackwell. doi:10.1002/9780470756492.ch6
- Sorace, Antonella. 2004. “Native Language Attrition and Developmental Instability at the Syntax-Discourse Interface: Data, Interpretations and Methods”. *Bilingualism: Language and Cognition* 7.143–145. doi:10.1017/S1366728904001543
- Sorace, Antonella. 2005. “Selective Optionality in Language Development”. *Syntax and Variation: Reconciling the Biological and the Social* ed. by Leonie Cornips & Karen P. Corrigan, 55–80. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/cilt.265.04sor
- Sorace, Antonella. 2011. “Pinning Down the Concept of ‘Interface’ in Bilingualism”. *Linguistic Approaches to Bilingualism* 1:1.1–33. doi:10.1075/lab.1.1.01sor
- Sorace, Antonella & Ludovica Serratrice. 2009. “Internal and External Interfaces in Bilingual Language Development: Beyond Structural Overlap”. *International Journal of Bilingualism* 13:2.195–210. doi:10.1177/1367006909339810
- Tsimpili, Ianthi & Antonella Sorace. 2006. “Differentiating Interfaces: L2 Performance in Syntax-Semantics and Syntax-Discourse Phenomena”. *Proceedings of the 30th Annual Boston University Conference on Language Development (BUCLD 30)* ed. by David Bamman, Tatiana Magnitskaia & Colleen Zaller, 653–664. Somerville, Mass.: Cascadilla.
- Tsimpili, Ianthi, Antonella Sorace, Caroline Heycock & Francesca Filiaci. 2004. “First Language Attrition and Syntactic Subjects: A Study of Greek and Italian Near-Native Speakers of English”. *International Journal of Bilingualism* 8.257–277. doi:10.1177/13670069040080030601
- White, Lydia. 1985. “The Pro-Drop Parameter in Adult Second Language Acquisition”. *Language Learning* 35:1.257–277. doi:10.1111/j.1467-1770.1985.tb01014.x
- White, Lydia. 1986. “Implications of Parametric Variation for Adult Second Language Acquisition: An Investigation of the ‘Pro-Drop’ Parameter”. *Experimental Approaches to Second Language Acquisition* ed. by Vivian Cook, 55–72. Oxford: Pergamon.
- White, Lydia. 2009. “Grammatical Theory: Interfaces and L2 Knowledge”. *The New Handbook of Second Language Acquisition* ed. by William C. Ritchie & Tej K. Bhatia, 49–68. Bingley, UK: Emerald Group.
- White, Lydia, Elena Valenzuela, Martyna Kozłowska-Macgregor & Yan Kit Ingrid Leung. 2004. “Gender and Number Agreement in Nonnative Spanish”. *Applied Psycholinguistics* 25.105–133. doi:10.1017/S0142716404001067



## PART II

# Internal and external factors in pragmatic variation



# Afro-Hispanic contact varieties at the syntax/pragmatics interface

## Pro-drop phenomena in Chinchano Spanish

Sandro Sessarego<sup>1,2,3,4</sup> & Javier Gutiérrez-Rexach<sup>5</sup>

<sup>1</sup>University of Texas at Austin / <sup>2</sup>Freiburg Institute for Advanced Studies /

<sup>3</sup>Helsinki Collegium for Advanced Studies / <sup>4</sup>Foro Latinoamericano de

Antropología del Derecho / <sup>5</sup>The Ohio State University

The Null Subject Parameter (NSP) has been the focus of much debate in the syntactic and pragmatic literature. Within the realm of Spanish and Portuguese, the analysis of two dialects that do not follow its predictions (Dominican Spanish (DS) and Brazilian Portuguese (BP)) has led to the postulation of new hypotheses to account for their unexpected syntactic patterns. The present study pays attention to yet another dialect of Spanish that does not conform to the NSP, Chinchano Spanish (CS), an Afro-Hispanic variety spoken in Chíncha, Peru. In so doing, this paper provides an analysis of null and overt subjects that partially deviates from previous accounts of similar pro-drop phenomena. Additionally, this study proposes a model of contact-induced language transmission that explains why CS – as well as many other Afro-Hispanic languages of the Americas (AHLAs) – presents patterns that do not align this dialect with either null-subject languages (NSLs) like Italian or non-null-subject languages (NNSLs) like English.

### 1. Introduction

This article focuses on a set of data at the syntax/pragmatics interface that have been related to the Null Subject Parameter (NSP) (Chomsky 1981; Rizzi 1982) and that can be found in Chinchano Spanish (CS), an Afro-Hispanic dialect spoken in the Province of Ica, coastal Peru: (a) the presence of non-emphatic, non-contrastive overt pronominal subjects; (b) an impoverished verb-agreement paradigm; and (c) the presence of non-inverted questions. The existence of such patterns in a number of Afro-Hispanic languages has traditionally been ascribed



to their supposed creole origin (cf. Perl & Schwegler 1998). The present analysis suggests that these structures can be accounted for as the result of conventionalized advanced SLA strategies, which do not necessarily imply any previous creole stage (cf. Sessarego 2013a). Additionally, a closer analysis of these phenomena in CS provides new evidence for the debate on the current status of the Null Subject Parameter (NSP), the validity of its original predictions – rooted in the Principles and Parameters (P&P) framework (Chomsky & Lasnik 1993) – its shortcomings (cf. Huang 1994; Holmberg 2005) and possible adjustments within the research agenda put forward by the Minimalist Program (MP) (Chomsky 1995, 2001, et seq.; Boeckx 2011).

This paper consists of six sections. Section 2 briefly summarizes the original P&P assumptions in relation to the NSP. Section 3 presents data from two Romance dialects that have generated much interest in recent years, since they do not appear to align either with null-subject languages (NSL) or with non-null subject ones (NNSL): Dominican Spanish (DS) and Brazilian Portuguese (BP). This section also offers an account of CS, another variety that acts as a semi-NSL. Section 4 proposes a novel, unified model to account for the data in DS, BP and CS. Section 5 consists of an account of contact-induced language transmission that may explain why CS, as well as several other Afro-Hispanic languages of the Americas (AHLAs), and potentially also DS and BP present the aforementioned syntactic features. Finally, Section 6 discusses our results and concludes the study.

## 2. The null-subject parameter (NSP)

Within the traditional Principles and Parameters (P&P) approach (Chomsky 1981; Chomsky & Lasnik 1993), languages can essentially be seen as the result of the interaction between the general principles of Universal Grammar (UG), of an innate nature and shared by all languages or linguistic varieties, and a finite set of binary parameters that are responsible for the syntactic variability observable across human languages. Within this framework, principles are “language-invariant statements” (Chomsky 1995: 25), whereas parameters must be set for certain values. Principles are part of an innate UG, which is part of our genetic endowment as members of the human species, i.e., the “language organ”. As such, they do not need to be learned by exposure to language since they are hypothesized to precede specific acquisition processes. Rather, exposure to language merely triggers the parameters to adopt a certain setting. Differential settings are associated with different languages (or varieties).

Chomsky (2000) uses an insightful metaphor when comparing the language faculty to a switch box or similar device. This box would consist of two components:

a fixed network, which would represent the innate principles of language, and several switches, which are options determined by experience, representing binary parameters that can be set on or off. Different parametric combinations lead to different grammars. In Chomsky's (2000: 8) words:

When the switches are set one way, we have Swahili; when they are set another way, we have Japanese. Each possible human language is identified as a particular setting of the switches – a setting of parameters, in technical terminology.

The NSP has played a major role in the research agenda of generative grammarians for at least three decades. In fact, starting with the early work by Rizzi (1982) on null and overt subjects, as incorporated in the canonical works on the P&P model (Chomsky 1981), this parameter has been taken as a good example of how apparently unrelated syntactic phenomena may be explained as the overt result of a single cluster of properties, all triggered by a particular setting of one 'switch'.

As recently summarized by Roberts & Holmberg (2010: 16–18), an NSL, a language with a value [+ ] for the NSP, was expected to show the cluster of properties presented in (1)–(4), while an NNSL, a language with the value [– ] for the NSP, was expected to lack them.

- (1) A silent, referential, definite subject of finite clauses
  - a. *Parla italiano* (Italian, +NSP)
  - b. \*Speaks Italian (English, –NSP)  
“She speaks Italian.”
- (2) Free subject inversion
  - a. *Hanno telefonato molti studenti* (Italian, +NSP)
  - b. \*Ont téléphoné beaucoup d'étudiants (French, –NSP)  
“Many students have telephoned.”
- (3) Absence of complementizer-trace effects
  - a. *Chi hai detto che – ha scritto questo libro?* (Italian, +NSP)
  - b. \*Who did you say that – wrote this book? (English, –NSP)  
“Who did you say wrote this book?”
- (4) Rich agreement inflection on finite verbs
  - a. *Yo como, tú comes*, etc. (Spanish, +NSP)
  - b. *I eat, you eat*, etc. (English, –NSP)

Subsequently, as pointed out by Camacho (2008, 2013), other properties were attributed to +NSP languages: expletives must be null (5), and overt pronouns cannot take an arbitrary reading (6) (Suñer 1983; Jaeggli 1986).

- (5) Expletives must be null
  - a. *Llueve.* (Spanish, +NSP)
  - b. *It rains.* (English, -NSP)
- (6) Overt pronouns cannot take an arbitrary reading
  - a. *Dijeron que habían venido* (Specific or arbitrary reading)  
 said-they that have-they come  
 "They said that they had come."
  - b. *Ellos dijeron que habían venido* (Specific reading)  
 they said that have-they come  
 "They said that they had come."

This latter generalization was formalized in Montalbetti's (1984) Overt Pronoun Constraint, which states that overt pronouns cannot be interpreted as a bound variable in a given language when such a language has null *pro*, which is the element used for this type of dependent reading (7) (cf. Camacho 2008: 417–418).<sup>1</sup>

- (7) Overt-Pronoun Constraint
  - a. *Todo estudiante<sub>i</sub> cree que pro<sub>i</sub> es inteligente*  
 every student thinks that pro is intelligent  
 "Every student thinks that he is intelligent."
  - b. *Todo estudiante<sub>i</sub> cree que él<sub>\*i/j</sub> es inteligente*  
 every student thinks that he is intelligent  
 "Every student thinks that he is intelligent."

The formulation of the NSP triggered a variety of empirical studies, which served as a testing ground for its predictions. The outcomes of such an effort resulted in the reformulation and adjustment of the original proposal (i.e., Jayaseelan 1999; Tomioka 2003; Saito 2004; Holmberg 2005) and, in some cases, also in its complete rejection (i.e., Newmeyer 2004; Haspelmath 2006). In fact, the study of a number of typologically different languages, in particular Asian languages (Vietnamese, Thai, Japanese, etc.), showed that the NSP predictions did not hold cross-linguistically. For example, in Chinese, a language without verbal agreement, pronouns can be null if certain discourse requirements are met (cf. Huang 1994).

---

1. A reviewer observes that there is a contrast regarding post-verbal pronouns that should somehow be noted: when the pronoun is post-verbal, no Montalbetti effect arises: *Todo estudiante cree que es inteligente él*. Why this is the case remains to be clarified. This is definitely something that deserves further attention since it will contribute to the analysis of SV/VS asymmetries.

It has been shown that even within the realm of Romance certain varieties present a mix of properties that appear to belong to both NSLs and NNSLs; thus their existence weakens the fundamental assumptions on which parametric syntax has been built in the sense of the binary character of parameters assumed by this theory. Such languages have often been labeled as “partial pro-drop systems”, and, with the advent of the Minimalist Program (MP), new hypotheses have been proposed within the generative enterprise to account for them. Two such varieties are Dominican Spanish and Brazilian Portuguese, and the following section presents some of the main claims that have been made to analyze pro-drop phenomena in these systems.

### 3. A look at two “partial pro-drop systems” in Romance

With the advent of the Minimalist Program (Chomsky 1995), the prevailing view of linguistic variation and contrast within Universal Grammar has undergone a shift from rigidly defined parameters – sometimes associated with clusters of properties (Chomsky 1981; Chomsky & Lasnik 1993) – to an approach in which features play a central role, are flexibly distributed and originate in the lexicon, according to what Baker (2008) calls the “Borer-Chomsky conjecture” (Borer 1984; Chomsky 2001). There is thus a reassessment of the locus of cross-linguistic variation: from innate parameters – coming with clusters of properties – to features more freely distributed across the lexicon (Borer 1984). This prompted the formulation of new hypotheses rooted in the more cautious observation of the empirical evidence. Dominican Spanish (DS) and Brazilian Portuguese (BP) are two dialects that have generated much interest in relation to the NSP predictions, since neither appears to conform to properties traditionally ascribed to either null-subject languages (NSLs) or non-null-subject ones (NNSLs).

Several studies have pointed out how non-emphatic, non-contrastive overt pronouns are extremely common in these languages to the point where many of the features traditionally differentiating null and overt pronouns in +NSP varieties appear to have been lost (Toribio 1993; Duarte 1993, 1995; Modesto 2000; Ordóñez & Olarrea 2006; among others). In addition to high rates of overt subjects, these languages also show a combination of properties that place them somehow in between NSLs and NNSLs. The halfway nature of these grammars has gained them the label of ‘partial pro-drop systems’. Some studies have tried to account for the variation attested in these dialects by postulating that speakers of these varieties are somehow bilingual in their own language, thus they would be able to switch between +NSP and –NSP configurations within a single

grammar (Toribio 2000), while others prefer to see the phenomena found in these dialects as the result of a change in progress, where NSL systems would be in the process of turning into NNSLs (Camacho 2008, 2013).

### 3.1 Dominican Spanish

Toribio (2000: 315–316) begins her well-known study on DS pro-drop phenomena by asserting a statement that is symptomatic of the long-lasting academic clash between syntacticians and sociolinguists:

Linguistic theorizing, as carried out within the Principles and Parameters framework, has proceeded, to date, largely uninformed by pertinent facts revealed within the sub-discipline of dialectology. While the theory has been significantly articulated by reference to proposals regarding attested morphosyntactic differences between languages (e.g., Spanish versus English, or Spanish versus French), the study of syntactic variation as observed between and within dialects of the same language (e.g., Standard Latin American Spanish versus Dominican Spanish), with notable exception, has been relegated to the domain of sociolinguistics.

Toribio then proceeds to present the idiosyncrasies of DS in relation to the NSP. As can be observed in the following examples, DS presents a mix of constructions that are not supposed to be possible within a single language, if the NSP predictions really apply. Toribio points out and gives examples (reproduced here) of the weakening of verbal agreement due to widespread phonological processes of consonant deletion (8), the coexistence of non-emphatic, non-contrastive overt subjects and null subjects (9) and the presence of overt expletive pronouns (10).

(8) Verbal agreement weakening

- a. Norm: [sal.'tar], [sal.'tas], [sal.'ta.β̃an]
- b. Santo Domingo: [sal.'tal], [sal.'ta], [sal.'ta.β̃an]
- c. Cibao Valley: [saj.'taj], [saj.'ta], [saj.'ta.β̃an]
- d. Southern coast: [sar.'tar], [sar.'ta], [sar.'ta.β̃an]

(9) Coexistence of non-emphatic, non-contrastive overt subjects and null subjects

- a. *Si ellos me dicen que yo estoy en peligro cuando ellos me entren la aguja por el ombligo, yo me voy a ver en una situación de estrés*  
 “If they tell me that I am in danger when they put the needle in my belly-button, I am going to find myself in a stressful situation.”
- b. *Yo no lo vi, él estaba en Massachusetts, acababa*  
 I not him saw, he was in Massachusetts finished  
*de llegar ...*  
 of arrive  
 “I did not see him, he was in Massachusetts, he just arrived ...”

- (10) Overt expletive pronoun
- Ello llegan guaguas hasta allá*  
it arrive buses to there  
“There arrive buses up to there.”
  - Ello había mucha gente*  
it was many people  
“There were a lot of people.”

Moreover, Toribio points out other phenomena that were not necessarily listed as properties deriving from a specific setting of the NSP but which appear to be relevant to the analysis since they concern the use of pronouns: (11) lack of S-V inversion in questions, (12) use of pseudo-cleft constructions as questions and (13) the use of copulative structures with a focalized constituent as a focus strategy (Bosque 1999; Camacho 2006; Méndez-Vallejo 2010).

- (11) Lack of S-V inversion in questions
- ¿Qué número tú anotaste?*  
what number you wrote-down  
“What number did you write down?”
  - ¿Qué yo les voy a mandar a esos muchachos?*  
what I them go to send to those boys  
“What am I going to send to those boys?”
- (12) Pseudo-cleft constructions
- ¿Dónde fue que tú estudiaste?*  
where was that you studied  
“Where did you study?”
  - ¿En qué es que tú te vas a graduar?*  
in what is that you REFL go to graduate  
“What will you graduate in?”
- (13) Copular focus strategy
- Yo quiero es comida*  
I want is food  
“I want food.”
  - Ese niño está es enfermo*  
that child is is sick  
“That child is sick.”

After reviewing these phenomena and the contexts in which they appear in DS, Toribio proposes an explanation for the observed patterns. She claims that DS is going through intensive grammatical restructuring, where linguistic parameters

would be in the process of re-setting (2000: 328). She adopts several minimalist tenets (Chomsky 1995), such as the distinction between weak and strong features and Pollock's (1989) Split-INFL Hypothesis, according to which tense and agreement information is checked under two different nodes. In particular, her analysis focuses on the competition of morphological strengths of nominal features on the functional nodes TENSE and AGR. Since the verbal features contained in AGR and TENSE are supposed to be strong in Spanish, the verb must move to such nodes. AGR and TENSE are also responsible for checking the phi-features and case features of subject NPs in their Spec positions. Given that strong phi-features in AGR are needed to license null subjects (cf. Chomsky 1995), Toribio assumes that Standard Latin American Spanish AGR is endowed with them. Moreover, since in this dialect VSO word order co-exists with SVO in declarative constructions, she proposes a structure in which AGR is syntactically dominated by TENSE.

On the other hand, according to Toribio's hypothesis, in Dominican Spanish, given the intense process of grammatical restructuring, AGR and TENSE may contain both strong and weak features. When AGR is endowed with weak nominal features, overt raising of NPs to Spec AGR would be prevented and – consequently – null subjects would not be allowed. In a similar fashion, when TENSE features are weak, only SVO constructions may occur, since VSO word order would be hampered.

Thus in DS it is possible to find the parameter setting represented by the properties of null subjects and postverbal subjects, as in Standard Latin American Spanish, as well as the innovative setting instantiated by lack of null subjects and preverbal subjects.

Building on preceding studies on diachronic syntax (i.e., Kroch 1991; Roberts 1993), Toribio characterizes DS as a synchronic case of co-existing competing grammars (Roeper 1999; Yang 2002). Thus DS speakers would have a variable I-language, containing competing parametric configurations. From her point of view, therefore, DS speakers would be “bilinguals in their native language, acquiring two grammars with opposed, competing values for the relevant parameters” (Toribio 2000: 339), and thus the reason behind the variation attested in the data may be conceived as an alternation – or code-switch – between the two available grammars.

### 3.2 Brazilian Portuguese

More recently, Camacho (2008, 2013) summarizes a variety of studies on Brazilian Portuguese pro-drop phenomena and points out patterns that do not follow the predictions of the NSP. In particular, building on Duarte's (1995) research, he

highlights the fact that the rate of overt pronouns in BP has increased steadily during the last century so that it went from 20% in 1845 to 74% in 1992. Such an increase in the use of overt pronouns parallels the erosion of the inflectional verb paradigm, as well as the reorganization of the pronominal system (see Tables 1 and 2) (Camacho 2008, adapted from Duarte 1995: 19; cf. also Kato & Negrão 2000).

**Table 1.** Evolution of inflectional paradigms in BP

Person/number	Paradigm 1	Paradigm 2	Paradigm 3
Speaker.SG	am-o	am-o	am-o
Hearer.SG	am-a-s	--	--
	am-a	am-a	am-a
Other.SG	am-a	am-a	am-a
Speaker.PL	am-a-mos	am-a-mos	--
	--	am-a	am-a
Addressee.PL	am-a-is	--	--
	am-a-m	am-a-m	am-a-m
Other.PL	am-a-m	am-a-m	am-a-m

**Table 2.** Evolution of the pronominal paradigm in BP

Person/number	Pronoun		Verbal ending
Speaker.SG	Eu	Eu	am-o
Adressee.SG	Tu	Você	am-a
	Vós		
Other.SG	Ele/Ela	Ele/Ela	am-a
Speaker.PL	Nós	A gente	am-a
Addressee.PL	Vós	Vocês	am-a
	Vocês		
Other.PL	Eles/Elas	Eles/Elas	am-a-m

Since such a reduced verbal inflection is no longer able to satisfy the EPP, an overt (weak) pronoun must move to Spec IP to meet such a requirement. This would allow the use of subject clitic left dislocations (SCLLDs), as in (14) (cf. Barbosa et al. 2005).

(14) SCLLDs

- a. *A Clarinhai elai cozinha que é uma maravilha*  
the Clarinha she cooks thay is a wonder  
‘Clarinha, she cooks wonderfully.’



- b. *Então [o Instituto de Física], elei manda os piores professores*  
 then the Institute of Physics he sends the worst professors  
*... [Os melhores], elesj dão aula no curso de matemática*  
 the best they give class in-the course of Mathematics  
 “Then the Institute of Physics sends the worst professors .... The best  
 teach Mathematics.”

Moreover, as is well known, BP lacks S-V inversion in questions, and according to Duarte & Kato (2002) this is a change that took place in the 19th century and correlates with the increase of overt pronouns (15).

(15) Lack of S-V inversion (Silva 2001)

- a. *O que a Maria leciona?*  
 what the Maria teaches  
 “What does Maria teach?”
- b. *\*O que leciona a Maria?*  
 what teaches the Maria
- c. *Onde ela leciona?*  
 where she teaches  
 “Where does she teach?”
- d. *\*Onde leciona ela?*  
 where teaches she

As in the case of DS, in BP null and overt pronouns co-exist, and depending on the context their interpretation may differ, as indicated by Modesto (2000: 152) and reported by Camacho (2013), as seen in (16).

(16) Coexistence of overt and null pronouns

- a. *O Pauloi convenceu o Pedroj que proi/\*j/\*k tinha que*  
 the Paulo convinced the Pedro that pro had that  
*ir embora.*  
 go away  
 “Paulo convinced Pedro that he had to go away.”
- b. *O Pauloi convenceu o Pedroj que elei/j/k tinha que*  
 the Paulo convinced the Pedro that pro had that  
*ir embora.*  
 go away  
 “Paulo convinced Pedro that he had to go away.”

Interestingly, the high rate of overt pronouns also questions the validity of Montalbetti’s (1984) Overt Pronoun Constraint, since in BP both overt and null pronouns

can be interpreted as bound variables (17) (cf. Barbosa et al. 2005: 43, in Camacho 2008: 422).

- (17) Overt Pronoun Constraint violation
- a. *[Ninguém no Brasil]<sub>i</sub> acha que ele<sub>i</sub> é prejudicado pelo Governo*
  - b. *[Ninguém no Brasil]<sub>i</sub> acha que pro<sub>i</sub> é prejudicado pelo Governo*  
 “No one in Brazil thinks that he is harmed by the government.”

After providing an overview of these phenomena for BP – and contrasting them with those found in DS – Camacho concludes that these dialects should be seen as grammars in change. In particular, the author suggests that these varieties are in the process of shifting from being NSL to being NNSLs. Camacho (2008: 426) formalizes two possible paths of change, which we schematically reproduce in (18). The locus of variation in Camacho’s proposal would be the lexicon.<sup>2</sup>

- (18) Two possible paths of change from NSL to NNSL
- a. Higher frequency → [+/-ref] pronominal → Pronominal in Spec, IP → Pronominal satisfies EPP, becomes weak, INFL becomes [-ref]
  - b. Higher frequency → Pronominal in Spec, IP → Pronominal satisfies EPP, becomes weak, INFL becomes [-ref] → pronominal [+/-ref]

### 3.3 Chinchano Spanish

We now turn our attention to a different variety of Spanish, Chinchano Spanish (CS), an Afro-Hispanic language spoken in Chincha, Peru, by the descendants of the slaves taken to this region to work in Jesuit sugarcane plantations during colonial times (cf. Sessarego 2014, 2015).<sup>3</sup> CS, in line with BP and DS, may be classified as a partial pro-drop system, in that structures that belong to NNSLs and NSLs are both attested. CS is a hybrid system, where overt subjects

---

2. Martínez-Sanz (2011) provides clear evidence for language change, but she identifies the change in progress as targeting the position of the subject, not so much in the null/overt subject contrast. This would entail that the idea of language change can be refined, and this may have consequences for the evaluation of Camacho’s proposal and the challenge for the NSP.

3. We first carried out sociolinguistic interviews and only in a second phase we asked for grammaticality judgments. This allowed us to compare speakers’ grammatical intuitions with naturalistic data. The result of such a comparison allowed us to get a more realistic picture of the Afro-Peruvian grammar and to avoid potential speech accommodations to what speakers may perceive as a more appropriate register.

are used redundantly, without signaling either emphasis or contrast, as in example (19).

(19) Non-emphatic, non-contrastive overt subjects

- a. *Mauricio fue también. Él se tomó una botella de cerveza*  
 Mauricio went too. He CL took a bottle of beer  
*y después él se fue al bar de fiesta*  
 and after he CL went to-the bar of party  
 “Mauricio went too. He drank a bottle of beer and afterwards he went to the bar to have fun.”
- b. *Cuando nosotros trabajamos, nosotros lo hacemos de verdad. Ellos*  
 when we work, we it do of real they  
*dicen que ellos trabajan pero ellos no trabajan duro*  
 say that they work but they not work hard  
 “When we work, we do it for real. They say that they work but they do not work hard.”

Additionally, CS shows variable subject-verb agreement (20). In fact, third-person singular forms may appear as default forms, especially in the speech of the oldest informants. In some cases, the lack of verb inflection may be prompted by phonological processes of word-final consonant weakening (20a, b) (cf. Sessarego 2015: Chapter 4 for an overview of CS phonetics and phonology) while, in other instances, the absence of verb inflection is clearly related to morphosyntax (20c, d).

(20) Lack of agreement

- a. *Ellas comía[n] lo que yo cocinaba*  
 they ate it that I cooked  
 “They ate what I cooked.”
- b. *¿Tú cree[s] que eso es así?*  
 you think that that is so  
 “Do you think that is like this?”
- c. *Yo compró un pedazo de tierra*  
 I bought a piece of land  
 “I bought a piece of land.”
- d. *Nosotros vivía con poca plata*  
 we lived with little silver  
 “We used to live with little money.”

Our data from this dialect clearly indicate the presence of questions where subject and verb have not been inverted (21a, b). Nevertheless, we wish to mention that not all questions present such a pattern; thus verb-inversion patterns are also attested in CS (21c, d).

## (21) Questions

- a. *¿Cómo uté se llama?*  
how you CL call  
“What is your name?”
- b. *¿Cuándo tú vuelve?*  
when you returning  
“When are you coming back?”
- c. *¿Qué quieres comer {tú}?*  
what want eat you.SG  
“What do you want to eat?”
- d. *¿Adónde van {ustedes}?*  
to-where go you.PL  
“Where are you going?”

CS also presents another property that has been ascribed to NNSL: overt pronouns can take an arbitrary reading (22) (cf. Suñer 1983; Jaeggli 1986). This also implies a potential violation of Montalbetti’s (1984) Overt Pronoun Constraint (23).

(22) *Overt pronouns can take an arbitrary reading*

- a. *Dijeron que habían venido* (Specific or arbitrary reading)  
said that have-they come  
“They said that they had come.”
- b. *Ellos? dijeron que habían venido* (Specific or arbitrary reading)  
they said that have-they come  
“They said that they had come.”

## (23) Overt-Pronoun-Constraint violation

- a. *Todo estudiantei cree que proi es inteligente*  
every student thinks that pro is intelligent
- b. *Todo estudiantei cree que éli/j es inteligente*  
every student thinks that he is intelligent  
“Every student thinks that he is intelligent.”

#### 4. The proposal: Differential specification of functional heads

When we take a close look at the CS data, the hypothesis of a constant code-switching (argued by Toribio among others for DS) appears to be counterintuitive for this dialect, since it is difficult to explain why speakers would keep code-switching in the absence of environments triggering such alternation. There is also lack of exposure to fully differentiated systems of a completely

different nature, as is the case for Spanish-English code-switchers. In other words, in CS there do not seem to be speakers in the community instantiating the separate languages triggering code-switching.<sup>4</sup> An analysis in terms of the 'grammars in change' idea (cf. Camacho's (2013) proposal above) is also insightful but needs supplementation. We do acknowledge that – to a certain extent – at the social/community level all grammars are in constant change, but at least at the idiolectal level, based on the grammaticality judgments of our informants, we would like to pursue an explanation where some stability is assumed – in line with mainstream generative assumptions. This line of analysis is consistent with Embick's (2008) claims that the question of whether there is a sociolinguistic effect on the distribution of variants can be kept distinct from the study of the constructions under consideration.

If we adopt one of the main tenets of the Chomskyan view of grammar, there is a fixed inventory of computational mechanisms that are universal (the Universal Grammar hypothesis) and change or variation take place at a 'micro' level of analysis. One of the leading assumptions of the Minimalist Program is that such computational devices or operations are minimal and constant in nature (internal and external MERGE, AGREE). Differential outputs in linearization are the result of contrasting feature specifications in the lexical items involved (at the initial starting point for the computation or 'numeration'). Stating this differently, the hypothesis that we want to advocate is that we should approach the phenomena associated with a hybrid NSP system such as CS not as an instance of code-switching or of competing or evolving grammars but as an instance of competing or evolving lexical and functional entries, *à la* Adger & Smith (2005). What this means is that, for a variety of reasons, the inventory or feature specification of lexical or functional elements differs not only across languages but also across varieties, making room for a more fine-grained treatment of sociolectal and idiolectal variation. More specifically, lexical or functional items with two possible specifications may coexist in a given variety. The choice of one or the other will have consequences in the word order arrangements that might become available. Therefore, the instances of variable subject-verb agreement found in this dialect can be accounted for by relying on the model proposed by Adger & Smith (2005) to formalize the variation between *was* and *were* forms in Buckie English. Specifically, we may envision a grammar in which two different Tense Heads (T) may be

---

4. Although it is possible that some speakers in this population have access to speakers of standard Peruvian Spanish (or to media in which this variety is predominant), this interaction does not exceed the normal amount of interaction attested in other varieties, which would not explain this degree of flux.

selected from the mental lexicon: T1 and T2. T1, like standard Spanish, would be endowed with tense, case, number and person features, while T2 would not carry number and person features. If we conceive variation in these terms, we can now see how the application of the operation AGREE between a subject pronoun and T1 will result in a fully conjugated verb form (as in examples (24)–(25)).

- (24) T1 [tense:present, ucase:nom, unum:, upers:] .... pronoun [num:pl, pers:1, case:nom] → T1[tense:present, num:pl, pers:1] .... pronoun [num:pl, pers:1, case:nom]

- (25) Spell-Out: *Nosotros bailamos*  
We.NOM dance.PRESENT.1PL

In this scenario, the merger of the tensed verb form and the pronoun and the application of AGREE derives the agreement configuration that is standard in Latin American and Peninsular Spanish. Adopting the probe-goal model of syntactic computation (Chomsky 2000), the T head acts as a probe with unvalued features. Such features get valued by co-valuation or unification (Pollard & Sag 1994) with the pronominal goal. Person and number features are valued by matching or co-valuation. case is the feature that sets the goal as an active match and satisfies the EPP feature of the probe (Chomsky 2001).<sup>5</sup>

Let us now look at the alternative ‘non-agreeing’ configuration. In this alternative scenario, the probe is not specified for person and number. Thus, it targets the goal to satisfy its EPP (case) feature. Nevertheless, there is no co-valuation with the person and number features of the goal (pronoun) because T2 is not specified for them. Thus the agreement operation will not target such features of the pronoun. The system leads to default (third person singular) values of the goal, so there is no overgeneration (cf. Gutiérrez-Rexach & Sessarego 2014).

- (26) T2[tense:present, ucase:nom] .... pronoun[num:pl, pers:1, ucase:] → T2[tense:present] .... pronoun[num:pl, pers:1, case:nom]

- (27) Spell-Out: *Nosotros baila*  
We.NOM dance.PRESENT.3SG

If we now analyze the position of subjects and verbs in interrogative environments, we can observe that in CS, in line with several other Afro-Hispanic dialects

---

5. Subject-oriented EPP is not always put together with case. Subject-oriented EPP can be a D-feature on T (Chomsky 1995) or it can be the need for a Specifier of T (Chomsky 2001). As such it is independent of case. Accounts that put together the EPP and case as two sides of the same coin in general try to eliminate the EPP in favor of obligatory case assignment; cf. Ortega-Santos (2008) for discussion of these EPP issues in Spanish.

and with Caribbean Spanish varieties, it is possible to encounter both structures in which the fronted *wh*-operator (*wh*-OP) is directly followed by a subject (*wh*-S-V), and cases in which the first element after the *wh*-OP is the verb (*wh*-V-S), as in example (28).

(28) Afro-Hispanic/ Caribbean varieties

- a. ¿*Qué tú comes*?  
what you eat  
“What do you eat?”
- b. ¿*Qué comes {tú}*?  
what eat you  
“What do you eat?”

These subject-verb configurations partially diverge from the possibilities found in Mainland Spanish varieties, where the *wh*-S-V structure is not allowed (29).

(29) Mainland Spanish

- a. \*¿*Qué tú comes*?  
what you eat  
“What do you eat?”
- b. ¿*Qué comes {tú}*?  
what eat you  
“What do you eat?”

The problem of subject inversion or lack thereof in questions has commanded much recent attention (summarized in Villa-García 2015), and we will not be analyzing it here. What is more relevant for our purposes is to focus on the connections with theories of Second Language Acquisition (SLA) (cf. Sessarego 2013a; Gutiérrez-Rexach & Sessarego 2014). A number of studies on formal syntax and SLA have tried to cast light on the mechanisms behind the acquisition of *wh*-fronting and S-V inversion in questions. Findings have suggested that target-like proficiency may be difficult to obtain for second language speakers, particularly when their L1 does not recur to such syntactic strategies (e.g., in Chinese, Korean and Japanese) (cf. Johnson & Newport 1989; Birdsong 1992; White 1992; Martohardjono & Gair 1993; White & Juffs 1998; among others).<sup>6</sup>

---

6. As a reviewer points out, other lines of research that might be argued to be connected to our discussion relate to the processing costs associated with long-distance dependencies; more specifically, the processing costs associated with questions with and without inversion (Goodall 2004). In turn, one could argue that the cases of lack of inversion are favored by a canonicity constraint as proposed by Bever (2006, 2008). This canonicity constraint is what underlies the EPP in English.

The fact that *wh*-S-V questions appear to be common among Spanish creoles (cf. Holm & Patrick 2007) has led some researchers to postulate a potential creole origin for the Spanish dialects currently presenting such constructions (e.g., Perl 1998). While, in theory, such a connection is not impossible, we should keep in mind that lack of S-V inversion in questions is a very common phenomenon in SLA varieties of Spanish and that it may be encountered even in very advanced interlanguages (Pienemann 1998, 2005). This suggests that *wh*-S-V questions should not necessarily be seen as a creole indicator. Moreover, it must be highlighted that, in the AHLAs and in the Spanish varieties spoken in the Caribbean, not only is it possible to find the *wh*-S-V order (cf. 28b), but *wh*-V-S constructions are common, and recent quantitative analyses have actually shown that they are the most frequent ones (cf. Gutiérrez-Bravo 2008: 227 for Dominican Spanish). Nevertheless, traditionally, linguistic studies have analyzed *wh*-S-V constructions as the Caribbean Spanish counterparts to Mainland *wh*-V-S structures. Gutiérrez-Bravo (2005, 2007, 2008) has repeatedly pointed out the importance of realizing that both patterns should not simply be seen as equivalent constructions in two different dialects; rather, they co-exist in Caribbean Spanish, and their alternation obeys precise pragmatic constraints. Therefore, according to Gutiérrez-Bravo, while *wh*-V-S and *wh*-S-V are both possible structures in Caribbean Spanish, in Mainland Spanish the latter option is not allowed. Its equivalent form would be (30), where the subject is a sentence topic displaced to the left-peripheral position.

- (30) *Tú ¿qué comes?*  
 you what eat  
 ‘What do you eat?’

Gutiérrez-Bravo (2005, 2008) formulates the following Interrogative Clause Condition to explain the EPP requirement associated with interrogative clauses: A clausal Extended Projection is interrogative if the head of the highest phrase in the Extended Projection bears the feature [Q]. Given this proposal, in an example like (29b), TP represents the highest projection. This allows *wh*-OP to land in [Spec, T], and T° to acquire a [Q] feature from Spec-Head agreement with *wh*-OP. This movement operation satisfies the EPP requirement, and consequently allows the subject to remain in its VP internal position (31). Conversely, in (29a), *wh*-OP moves to [Spec, C], and C° acquires its [Q] feature. In such a configuration, a topicalized subject will have to land on [Spec, T] to satisfy the EPP (32).

- (31) [<sub>TP</sub> *Qué<sub>i</sub> comes* [<sub>VP</sub> *tú t<sub>j</sub> t<sub>i</sub>*]]?  
           wh T°  
           [Q] → [Q]



- (32)  $[_{CP} \text{Qué}_i \emptyset [_{TP} \text{tú}_j \text{comes}_k [_{VP} \text{t}_j \text{t}_k \text{t}_i]]]?$   
        $\text{wh C}^\circ$   
        $[Q] \rightarrow [Q]$

Following the reasoning pursued above, the proposal that we have advocated on the existence of two different functional heads (T1 and T2) with contrasting activity and agreement requirements can be extended to account for this asymmetry. When T1 enters the numeration and participates in the syntactic computation, it can target the Q feature of the *wh*-expression and satisfy its EPP requirement. Consequently, the subject remains in the *vP/VP* sphere and the *wh*-V-S order results. On the other hand, T2 is more impoverished in this respect and lacks the capacity to satisfy the EPP with Q. Thus, the Q feature of the *wh*-expression is co-valued (and agrees) with the C head (namely, it satisfies the EPP feature of C). Since the EPP feature of T2 is not checked yet, this requirement must be satisfied by the pronoun. The resulting spell-out is *wh*-S-V as expected. This account relies heavily on the role of the EPP. The availability of two options in Caribbean Spanish, namely the fact that it is optional to have the pronoun after the *wh*-item and before the verb in Caribbean, does not mean that satisfying the EPP is optional. It means that T1 and T2 do so differently. The issue of the connection with the discourse function of intercalated pronouns (Gutiérrez-Bravo 2005, 2008) remains a question for further research, since the EPP is just a formal feature, not discourse-oriented. Similarly, the absence of Montalbetti's effects seems to point in the direction of the critical role of T1/T2 differentiation and the formal role of EPP rather than to the association of subject pronouns with discourse features, although we discuss this aspect in more detail in the next section.

The existence in the AHLAs of two different word orders to express two different types of questions may be perceived as an additional complexity, not in line with the vast majority of contact linguistic phenomena, which tend to favor the acquisition of less complex/unmarked structures. This is definitely an issue that deserves to be studied in more depth. It requires a sociolinguistic analysis of the diachronic evolution of the *wh*-S-V order in the varieties presenting such a structure. Nevertheless, for the moment, a highly speculative proposal might suggest that at a certain point in the development of these dialects, due to processability constraints on L2 production (cf. Goodall 2004; Pienemann 2005), the linguistic output of a certain generation of L2 speakers may have been quite variable (including both inverted and non-inverted questions). Given such a context, it is not completely unreasonable to think that two different interpretations might have been assigned to such constructions by the acquiring children, so that in their L1 example (28a) came to represent the topicalized subject question that would be normally expressed with (30) in other Spanish dialects.

## 5. On the origin of Chinchano Spanish partial pro-drop system

After presenting this hypothesis, which may offer a unified account for the three dialects (BP, DS and CS), we provide an account of why CS, the vast majority of the AHLAs and potentially also DS and BP present partial pro-drop grammars. The use of non-emphatic, non-contrastive overt subjects is a linguistic phenomenon related to the acquisition of the null-subject parameter (Chomsky 1981; Rizzi 1982). Subject expression in null-subject languages such as Italian or Peninsular Spanish requires a certain degree of maturation in the acquisition of several constraints operating at the syntax/pragmatics interface, since both structural and discourse-related features appear to be at play. As has been observed, the null subject (*pro*) tends to be used in topic and non-contrastive focus contexts. The example in (33) illustrates the use of *pro* in Spanish (Montrul et al. 2009: 303). In this case, *pro* expresses old information, in the sense that pronominal reference can be resolved from a set of discourse entities already present in discourse.

- (33) *Juan llegó a su casa del trabajo. Primero pro se cambió*  
 Juan arrived to his home of-the work. First pro CL changed  
*de ropa y luego pro decidió ponerse a preparar la cena*  
 of cloth and then pro decided put-CL to prepare the dinner  
 “Juan came home from work. First he changed his clothes and then he  
 decided to make dinner.”

Recent investigations in SLA have repeatedly shown that even very advanced speakers tend to present difficulties in this aspect of L2 acquisition, thus resulting in the overproduction of overt subjects (cf. Montrul 2004 for a general review). The explanation would be that topic features are complex to acquire and, consequently, a native-like use of overt and covert pronouns is not likely to obtain (Sorace 2000, 2003, 2004). Grimshaw & Samek-Lodovici (1998) suggest that in pro-drop languages there is a main distinction between overt and covert subjects. This difference consists of the specification in the former ones of a [+topic shift] feature, which would be absent in the latter grammatical elements. On the other hand, such a different specification is not found in non-pro-drop varieties, where covert subjects (*pro*) do not exist. This proposal is obviously in line with the idea that certain aspects of the AHLAs are rooted in advanced L2 strategies (Sessarego 2013a). The grammatically adequate and pragmatically felicitous use of *pro* in Spanish implies the simultaneous knowledge of interrelating syntactic and pragmatic properties. Consequently, the fact that non-emphatic, non-contrastive overt pronominal subjects are consistently attested in these languages would not be surprising, since they can be viewed as instantiations of an advanced L2 competence.

Another significant issue in the analysis of the acquisition of Spanish as a second language is the host of phenomena derived from the relative mastering of phi-features (person, gender, number) (cf. Ramírez-Cruz, this volume). The morphological expression of these features can be mostly considered 'redundant' in this language and does not contribute to semantic interpretation. This can be explained as a byproduct of the 'non-interpretable' nature of gender in the Minimalist Program (Chomsky 1995). The complete mastery of such a pattern of morphological expression occurs relatively late in L2 acquisition and oftentimes is not achieved at all (Franceschina 2002). With respect to Spanish L2 grammars, the slow acquisition of agreement features results in Spanish interlanguages instantiating varying degrees of morphological incompleteness or departure from the standard variety in the marking of regularly agreeing expressions across the nominal and verbal domains. This property can be related to the well-attested observation that Spanish verb forms with invariant person and number are common in Spanish interlanguages and also in child language (Bybee 1985). In such cases, the most common pattern is the use of the third-person singular form as the default one. The CS variety and the other AHLAs display variable levels of subject-verb (dis)agreement. These phenomena, in turn, reflect an aspect of their degree of restructuring (Sessarego 2012): in some varieties, third-person singular default forms can be attested regularly (e.g., Afro-Bolivian Spanish, cf. Lipski 2008; Sessarego 2011) while, in other varieties, they are rare (e.g., Chota Valley Spanish, cf. Sessarego 2013b).

The proposal that we would like to present here is that CS, in line with several other AHLAs, can be conceived as THE RESULT OF L1 ACQUISITION (NATIVIZATION) OF ADVANCED L2 GRAMMARS (cf. Sessarego 2013a; Romero & Sessarego, this volume). By stating this, we assume that UG is responsible for the acquisition of both L1 and L2. Nevertheless, these two processes present significant differences. During L1 acquisition, the evolution of language proceeds naturally and instinctively, provided that the child is exposed to enough linguistic input. Conversely, L2 acquisition faces more challenges, since both biological and social factors appear to inhibit the full mastery of the TL. Indeed, biological age maturation corresponds to a progressive loss of spontaneity in acquisition, while certain social aspects of adulthood, such as lack of motivation, acculturation and free time may further conspire against a target-like attainment of the L2 (Herschensohn 2000: Chapter 3).

Given these assumptions, we wish to propose that CS, as well as many other AHLAs, may be seen as the result of a process of L1 acquisition (nativization) of advanced L2 grammars. This implies that, at a certain point in the evolution of these varieties, the enslaved Africans taken to the Americas had relatively good access to Spanish (the TL). Thanks to UG, they were able to acquire an advanced L2

variety of this language (G1). Their linguistic outputs (X) represented the primary linguistic data (PLD<sub>x</sub>) for their offspring, who acquired this language natively. The result of this model is a native variety of Spanish (G2), which has been built on L2 inputs. This accounts, therefore, for the presence of crystallized aspects of an L2 that can be systematically found across several L1s (in this case CS and the other AHLAs). This model is schematically represented in (34):<sup>7</sup>

- (34) a. Individual from Generation 1:  
       TLy → UG driving L2 acquisition → G1 → set of outputs X  
       b. Individual from Generation 2:  
       PLD<sub>x</sub> → L2 driving L1 acquisition → G2 → set of outputs Z

## 6. Conclusion

In this paper, we have presented an analysis of dialects in which a mixed system is attested with respect to the NSP. These dialects exhibit partial pro-drop grammars, in that instances of pro-drop phenomena and the usual constellation of grammatical features can be attested. Nevertheless, the standard arrangement of properties characteristic of NNSLs is also attested. An account of these facts is problematic for a standard view of parameters based on the idea/metaphor of a switchboard with several on/off switches. More recent accounts in terms of hypothesizing changing or competing grammars, although representing an improvement over the fixed-parameters idea, are somewhat inadequate to characterize what is going on in CS and potentially also in other mixed dialects. Thus, we argue that the mastery of overt and null pronouns, verbal agreement and inverted questions depends on advanced acquisition strategies which appear to be hampered by processability and language interface constraints (e.g., syntax/pragmatics and syntax/morphology interfaces). In line with Minimalist constructionist assumptions (Herschensohn 2000), the acquisition of the lexicon and of its formal features (Borer 1984) is supposed to develop gradually through a UG-driven path, and the differential specification of certain functional elements (in our case T) seems to play a pivotal role. This process generates several possible L2 grammars, with different parametric configurations. The variable second-language learners' output

---

7. Cf. Pires & Thomason (2008) and Pires & Rothman (2009) for a similar account of cross-generational language change. The main difference between their accounts and the present one is that example (34) pictures a case of contact-induced change, where SLA processes are involved. See also Veenstra (2008: 234–235) for a similar yet different account of nativization in creole genesis.

resulting from this acquisition process represents the PLD of the following generation, which will nativize the language into a new grammar. In general, this paper also provides evidence for arguments questioning the validity of the NSP or, more broadly, for recent proposals that revisit the concept of ‘parameter’ and suggest new potential paths of analysis (cf. Boeckx 2011).

## References

- Adger, David & Jennifer Smith. 2005. “Variation and the Minimalist Program”. *Syntax and Variation: Reconciling the Biological and the Social* ed. by Leonie Cornips & Karen P. Corrigan, 149–178. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/cilt.265.10adg
- Baker, Mark. 2008. *The Syntax of Agreement and Concord*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511619830
- Barbosa, Pilar, Mary A. Kato & Maria E. Duarte. 2005. “Null Subjects in European and Brazilian Portuguese”. *Journal of Portuguese Linguistics* 4.11–52.
- Bever, Tom G. 2006. “Minimalist Behaviorism: The Role of the Individual in Explaining Language Universals”. *Language Universals* ed. by Morten Christiansen, Christopher Collins & Shimon Edelman, 270–298. Oxford: Oxford University Press.
- Bever, Tom G. 2008. “Remarks on the Individual Basis for Linguistic Structures”. Ms., University of Arizona.
- Birdsong, David. 1992. “Ultimate Attainment in Second Language Acquisition”. *Language* 68. 706–755. doi:10.1353/lan.1992.0035
- Boeckx, Cedric. 2011. “Approaching Parameters from Below”. *The Biolinguistics Enterprise: New Perspectives on the Evolution of the Human Language Faculty* ed. by Anna Maria di Sciullo & Cedric Boeckx, 205–221. Oxford: Oxford University Press.
- Borer, Hagit. 1984. *Parametric Syntax. Case Studies in Semitic and Romance Languages*. Dordrecht: Foris.
- Bosque, Ignacio. 1999. “On Focus vs. Wh-Movement: The Case of Caribbean Spanish”. *Sophia Linguistica* 44–45.1–32
- Bybee, Joan. 1985. *Morphology: A Study of the Relation between Meaning and Form*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/tsl.9
- Camacho, José. 2006. “In Situ Focus in Caribbean Spanish”. *Selected Proceedings of the 9th Hispanic Linguistics Symposium* ed. by Nuria Sagarra & Almeida Jacqueline Toribio, 13–23. Somerville, Mass.: Cascadia.
- Camacho, José. 2008. “Syntactic Variation: The Case of Spanish and Portuguese Subjects”. *Studies in Hispanic and Lusophone Linguistics* 1:2415–433. doi:10.1515/shll-2008-1027
- Camacho, José. 2013. *Null Subjects*. Cambridge: Cambridge University Press. doi:10.1017/CBO9781139524407
- Chomsky, Noam & Howard Lasnik. 1993. “The Theory of Principles and Parameters”. *Syntax: An International Handbook of Contemporary Research*, vol. 1, ed. by Joachim Jacobs, Arnim von Stechow, Wolfgang Sternefeld & Theo Vennemann, 506–569. Berlin: Walter de Gruyter.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, Mass.: MIT Press.

- Chomsky, Noam. 2000. "Minimalist Inquiries: The Framework". *Step by Step* ed. by Roger Martin, David Michaels & Juan Uriagereka, 89–156. Cambridge, Mass: MIT Press.
- Chomsky, Noam. 2001. "Derivation by Phase". *Ken Hale: A Life in Language* ed. by Michael Kenstowicz, 1–52. Cambridge, Mass: MIT Press.
- Duarte, Maria E. 1993. "Do pronome nulo ao pronome pleno: A trajetória do sujeito no português do Brasil". *Português brasileiro: Uma viagem diacrônica* ed. by Ian Roberts & Mary A. Kato, 107–128. Campinas: Ed. Da Unicamp.
- Duarte, Maria E. 1995. *A perda do princípio "Evite pronome" no português brasileiro*. Ph.D. dissertation, UNICAMP.
- Duarte, Maria E. & Mary Kato. 2002. "A Diachronic Analysis of Brazilian Portuguese Wh-Questions". *Santa Barbara Portuguese Studies* 6.326–340.
- Embick, David. 2008. "Variation and Morphosyntactic Theory: Competition Fractioned". *Language and Linguistics Compass* 2.59–78. doi:10.1111/j.1749-818X.2007.00038.x
- Franceschina, Florencia. 2002. "Case and  $\phi$ -Feature Agreement in Advanced L2 Spanish Grammars". *EUROSLA Yearbook* ed. by Susan H. Foster-Cohen, Tanja Ruthenberg & Marie Louise Poschen, 71–86. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/eurosla.2.07fra
- Goodall, Grant. 2004. "On the Syntax and Processing of Wh-questions in Spanish". *Proceedings of the 23rd West Coast Conference on Formal Linguistics* ed. by Benjamin Schmeiser, Vineeta Chand, Ann Kelleher & Angelo Rodriguez, 237–250. Somerville, Mass.: Cascadilla.
- Grimshaw, Jane & Vieri Samek-Lodovici. 1998. "Optional Subjects and Subject Universals". *Is the Best Good Enough? Optimality and Competition in Syntax* ed. by Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky, 192–219. Cambridge, Mass.: MIT Press.
- Gutiérrez-Bravo, Rodrigo. 2005. *Structural Markedness and Syntactic Structure*. London & New York: Routledge/Taylor & Francis.
- Gutiérrez-Bravo, Rodrigo. 2007. "Prominence Scales and Unmarked Word Order in Spanish". *Natural Language and Linguistic Theory* 25.235–271. doi:10.1007/s11049-006-9012-7
- Gutiérrez-Bravo, Rodrigo. 2008. "Topicalization and Preverbal Subjects in Spanish Wh-Interrogatives". *Selected Proceedings of the 10th Hispanic Linguistics Symposium* ed. by Joyce Bruhn de Garavito & Elena Valenzuela, 225–236. Somerville, Mass.: Cascadilla.
- Gutiérrez-Rexach, Javier & Sandro Sessarego. 2014. "Morphosyntactic Variation and Gender Agreement in Three Afro-Andean Dialects". *Lingua* 151.142–161. doi:10.1016/j.lingua.2014.04.011
- Haspelmath, Martin. 2006. "Parametric versus Functional Explanations of Syntactic Universals". *The Limits of Syntactic Variation* ed. by Teresa Biberauer & Anders Holmberg. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/la.132.04has
- Herschenson, Julia. 2000. *The Second Time Around: Minimalism and L2 Acquisition*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/lald.21
- Holm, John A. & Peter L. Patrick, eds. 2007. *Comparative Creole Syntax: Parallel Outlines of 18 Creole Grammars*. London: Battlebridge.
- Holmberg, Anders. 2005. "Is There a Little *pro*? Evidence from Finnish". *Linguistic Inquiry* 36.533–564. doi:10.1162/002438905774464322
- Huang, Yan. 1994. *The Syntax and Pragmatics of Anaphora: A Study with Special Reference to Chinese*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511554292
- Jaeggli, Oswaldo. 1986. "Arbitrary Plural Pronouns". *Natural Language and Linguistic Theory* 4.43–76. doi:10.1007/BF00136264



- Jayaseelan, K. A. 1999. "Empty Pronoun in Dravidian". *Parametric Studies in Malayalam Syntax* ed. by K.A. Jayaseelan, 14–25. New Delhi: Allied Publishers.
- Johnson, Jacqueline S. & Elissa L. Newport. 1989. "Critical Period Effects in Second Language Learning: The Influence of Maturational State on Acquisition of English as a Second Language". *Cognitive Psychology* 21.60–99. doi:10.1016/0010-0285(89)90003-0
- Kato, Mary & Esmeralda Negrão, eds. 2000. *Brazilian Portuguese and the Null Subject Parameter*. Frankfurt & Madrid: Iberoamericana/Vervuert.
- Kroch, Anthony. 1991. "Reflexes of Grammar in Patterns of Language Change". *Language Variation and Change* 1.199–244. doi:10.1017/S0954394500000168
- Lipski, John. 2008. *Afro-Bolivian Spanish*. Madrid & Frankfurt: Iberoamericana/Vervuert.
- Martínez-Sanz, Cristina. 2011. *Null and Overt Subjects in a Variable System: The Case of Dominican Spanish*. Ph.D. dissertation, University of Ottawa.
- Martohardjono, Gita & James W. Gair. 1993. "Apparent UG Inaccessibility in SLA: Misapplied Principles or Principled Misapplications?" *Confluence: Linguistics, Second Language Acquisition and Speech Pathology* ed. by Fred R. Eckman, 79–103. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/lald.4.08mar
- Méndez-Vallejo, Catalina. 2010. "Syntactic Variation in Colombian Spanish. The Case of the Focalizing Ser (FS) Structure". *Romance Linguistics 2009: Selected Papers from the 39th Linguistic Symposium on Romance Languages (LSRL)* ed. by Sonia Colina, Antxon Olarrea & Ana Maria Carvalho, 169–186. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/cilt.315.10men
- Modesto, Marcello. 2000. "Null Subject without 'Rich' Agreement". In Kato & Negrão (2000), 147–174.
- Montalbetti, Mario. 1984. *After Binding: On the Interpretation of Pronouns*. Ph.D. dissertation, Massachusetts Institute of Technology.
- Montrul, Silvina. 2004. *The Acquisition of Spanish*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/lald.37
- Montrul, Silvina. Rejean Dias & Ana Thomé-Williams. 2009. "Subject Expression in the Non-Native Acquisition of Brazilian Portuguese". *Minimalist Inquiries into Child and Adult Language Acquisition: Case Studies across Portuguese* ed. by Acrisio Pires & Jason Rothman, 301–325. Berlin: Mouton De Gruyter. doi:10.1515/9783110215359.2.301
- Newmeyer, Frederick. 2004. "Against a Parameter-Setting Approach to Language Variation". *Linguistic Variation Yearbook* 4.181–234. doi:10.1075/livy.4.06new
- Ordóñez, Francisco & Antxon Olarrea. 2006. "Microvariation in Caribbean/Non-Caribbean Spanish Interrogatives". *Probus* 18.59–96. doi:10.1515/PROBUS.2006.003
- Ortega-Santos, Iván. 2008. *Projecting Subjects in Spanish and English*. Ph.D. dissertation, University of Maryland.
- Perl, Matthias. 1998. "Introduction". *América negra: Panorámica actual de los estudios lingüísticos sobre variedades hispanas, portuguesas y criollas* ed. by Matthias Perl & Armin Schwegler, 1–24. Madrid & Frankfurt: Iberoamericana/Vervuert.
- Perl, Matthias & Armin Schwegler, eds. 1998. *América negra: Panorámica actual de los estudios lingüísticos sobre variedades hispanas, portuguesas y criollas*. Madrid & Frankfurt: Iberoamericana/Vervuert.
- Pienemann, Manfred. 1998. *Language Processing and Second Language Development: Processability Theory*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/sibil.15
- Pienemann, Manfred, ed. 2005. *Cross-linguistic Aspects of Processability Theory*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/sibil.30

- Pires, Acrisio & Jason Rothman. 2009. "Introduction". *Minimalist Inquiries into Child and Adult Language Acquisition: Case Studies across Portuguese* ed. by Acrisio Pires & Jason Rothman, 5–34. Berlin: Mouton de Gruyter. doi:10.1515/9783110215359.0.5
- Pires, Acrisio & Sara G. Thomason. 2008. "How Much Syntactic Reconstruction Is Possible?". *Principles of Syntactic Reconstruction* ed. by Gisella Ferraresi & Maria Goldbach, 27–72. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/cilt.302.04pir
- Pollard, Carl & Ivan Sag. 1994. *Head-Driven Phrase Structure Grammar*. Chicago: The University of Chicago.
- Pollock, Jean-Yves. 1989. "Verb Movement, Universal Grammar, and the Structure of IP". *Linguistic Inquiry* 20.365–424.
- Rizzi, Luigi. 1982. *Issues in Italian Syntax*. Dordrecht: Foris. doi:10.1515/9783110883718
- Roberts, Ian. 1993. *Verbs and Diachronic Syntax*. Dordrecht: Kluwer.
- Roberts, Ian & Anders Holmberg. 2010. "Introduction: Parameters in Minimalist Theory". *Null Subjects: The structure of parametric variation* ed. by Teresa Biberauer, Anders Holmberg, Ian Roberts & Michael Sheehan, 1–57. Cambridge: Cambridge University Press.
- Roeper, Tom. 1999. "Universal Bilingualism". *Bilingualism: Language and Cognition* 2.169–184. doi:10.1017/S1366728999000310
- Saito, Mamoru. 2004. "Ellipsis and Pronominal Reference in Japanese Clefts". *Studies in Modern Grammar* 36.1–44.
- Sessarego, Sandro. 2011. *Introducción al idioma afroboliviano: Una conversación con el awicho Manuel Barra*. Cochabamba & La Paz: Plural Editores.
- Sessarego, Sandro. 2012. "Non-Creole Features in the Verb System of Afro-Hispanic Languages: New Insights from SLA Studies". *International Journal of Linguistics* 4.146–157. doi:10.5296/ijl.v4i1.1415
- Sessarego, Sandro. 2013a. "Afro-Hispanic Contact Varieties as Conventionalized Advanced Second Languages". *Iberia* 5.96–122.
- Sessarego, Sandro. 2013b. *Chota Valley Spanish*. Madrid & Frankfurt: Iberoamericana/Vervuert.
- Sessarego, Sandro. 2014. "Afro-Peruvian Spanish in the Context of Spanish Creole Genesis". *Spanish in Context* 11:3381–401. doi:10.1075/sic.11.3.04ses
- Sessarego, Sandro. 2015. *Afro-Peruvian Spanish: Casting Light on the Legacy of a "Missing Spanish Creole"*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/cll.51
- Silva, Glaucia. 2001. *Word Order in Brazilian Portuguese*. Berlin: De Gruyter. doi:10.1515/9783110869552
- Sorace, Antonella. 2000. "Syntactic Optionality in Non-Native Grammars". *Second Language Research* 16.93–102. doi:10.1191/026765800670666032
- Sorace, Antonella. 2003. "Near-Nativeness". *The Handbook of Second Language Acquisition* ed. by Catherine Doughty & Michael Long, 130–153. Oxford: Blackwell. doi:10.1002/9780470756492.ch6
- Sorace, Anonella. 2004. "Native Language Attrition and Developmental Instability at the Syntax-Discourse Interface: Data, Interpretations and Methods". *Bilingualism: Language and Cognition* 7.143–145. doi:10.1017/S1366728904001543
- Suñer, Margarita. 1983. "proarb". *Linguistic Inquiry* 14.188–191.
- Tomioka, Satoshi. 2003. "The Semantics of Japanese Null Pronouns and Its Cross-Linguistic Implications". *The Interfaces. Deriving and Interpreting Omitted Structures* ed. by Kerstin Schwabe & Susanne Winkler. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/la.61.16tom
- Toribio, Almeida J. 1993. "Lexical Subjects in Finite and Non-Finite Clauses". *Cornell Working Papers in Linguistics* 11.149–178.



- Toribio, Almeida J. 2000. "Setting Parametric Limits on Dialectal Variation in Spanish". *Lingua* 110.315–341. doi:10.1016/S0024-3841(99)00044-3
- Veenstra, Tonjes. 2008. "Creole Genesis: The Impact of the Language Bioprogram Hypothesis". *The Handbook of Pidgin and Creole Studies* ed. by Silvia Kouwenberg & John Singler, 219–241. Oxford: Blackwell.
- Villa-García, Julio. 2015. *The Syntax of Multiple-que Sentences in Spanish: Along the Left Periphery*. Amsterdam & Philadelphia: John Benjamins.
- White, Lydia. 1992. "Subacency Violations and Empty Categories in L2 Acquisition". *Island Constraints* ed. by Helen Goodluck & Michael Rochemont, 445–464. Dordrecht: Kluwer. doi:10.1007/978-94-017-1980-3\_17
- White, Lydia & Alan Juffs. 1998. "Constraints on *Wh*-Movement in Two Different Contexts of Non-Native Language Acquisition: Competence and Processing". *The Generative Study of Second Language Acquisition* ed. by Suzanne Flynn, Gita Martohardjono & Wayne O'Neil, 111–129. Mahwah, New Jersey: Lawrence Erlbaum.
- Yang, Charles. 2002. *Knowledge and Learning in Natural Language*. Oxford: Oxford University Press.

## CHAPTER 5

# Borrowed Spanish discourse markers in narrative

## A comparison across three generations of Tojol-ab'al (Mayan) speakers

Mary Jill Brody

Louisiana State University

This pilot study examines texts from three generations of speakers of Tojol-ab'al, a Mayan language that has been in long-term contact with Spanish, to determine the influence from Spanish on Tojol-ab'al narrative discourse. Discourse markers borrowed from Spanish and those indigenous to Tojol-ab'al are examined to determine discourse structure. Discourse markers borrowed from Spanish appear frequently even in the speech of the oldest, monolingual generation. Although the youngest generation uses many more words (nouns, verbs) borrowed from Spanish than do the earlier generations, narrative discourse structure, as indicated by the use of both borrowed and indigenous discourse markers, remains substantially the same through time.

### 1. Introduction

In this paper I report on a pilot study that demonstrates that the narrative discourse structure of Tojol-ab'al (a Mayan language spoken by several thousand people in a relatively remote region of Chiapas, Mexico) has remained substantially the same for three generations. To accomplish this, I compare the use of discourse markers (DMs) by speakers from three consecutive generations; I look at the use of both the indigenous DMs and those borrowed from Spanish into Tojol-ab'al. DMs are classically defined as “sequentially dependent elements which bracket units of talk” (Schiffrin 1987: 31). While this definition has been subsequently refined (see Sinnott, this volume; see also reviews of definitions in Brody 2010 and Matras 1998), it has the advantage of highlighting the role that these high-frequency elements play in discourse. The role of DMs in spoken

Spanish has been well defined (Cepeda & Poblete 1996; Poblete 1998; among others); DMs have also been analyzed for spoken Tojol-ab'al (Brody 1987, 1989, 1995, 1998; see Brody 2010 for a detailed comparison of Spanish and Tojol-ab'al DMs). Tojol-ab'al and Spanish are of course not genetically related, but they have been in contact for at least 450 years (Lenkersdorf 1986; for other discussion of Spanish in contact, see Philip, Romero & Sessarego; Sessarego & Gutiérrez-Rexach; and Kyzar, Smith & Schroth, all this volume). In contrast to the findings of earlier scholars (Haugen 1950), borrowing of DMs is a common feature of language contact; in fact, Matras identifies DMs (he calls them "utterance modifiers") as being "at the very top of the borrowability hierarchy in situations of conventional, interactional language contact" (1998: 282). Borrowing of DMs from Spanish occurs commonly in indigenous languages in contact with Spanish (Brody 1987, 1995; Stolz & Stolz 1996; Thomason & Kaufman 1988; Torres 2006), and DM marker borrowing is frequent in language contact in general (Matras 1998, 2006). Some investigators of borrowed DMs, e.g., Clyne (2003) and Matras (1998), have claimed that the borrowing of DMs changes the discourse structure of the borrowing language to more closely resemble that of the lending language. My findings here dispute that claim for the Tojol-ab'al data. Despite the long period of contact, many features of Tojol-ab'al discourse are not shared with Spanish, including a very high frequency of repetition, the use of overlay narrative structure (Grimes 1972), the impetus for dialogue (see Brody 2010) and the existence of an extensive set of indigenous DMs. I find for Tojol-ab'al that, while the borrowing of other categories of Spanish words shows increase over time, notably in the speech of the youngest generation, the structure of the discourse, as indicated by the consistent use of both borrowed and indigenous DMs, remains the same through time.

There is no question that the Spanish DMs are borrowed into Tojol-ab'al; they do not represent code-switches because they are found in the speech of monolinguals. Among the eldest generation in the pilot study, the speakers were monolingual in Tojol-ab'al, with only the smallest degree of passive bilingualism; that is, they understood a little Spanish but not much, and spoke only Tojol-ab'al. Even in the speech of bilinguals, the borrowed DMs do not trigger a switch into Spanish in the rest of the utterance. In addition, as mentioned above, there is a robust set of indigenous DMs. As I have shown elsewhere, the Tojol-ab'al proclivity for repetition is served well by the borrowing of Spanish DMs, which frequently occur along with the indigenous markers that have the same meaning (Brody 2010).

DMs have been examined in the context of language contact by a number of scholars (e.g., Brody 1987, 1989, 1998, 2010; Clyne 2003; Matras 1998, 2006;

Stolz & Stolz 1996; Torres 2006). In this way, DMs have been at the forefront of the incipient discipline of historical discourse analysis. Brinton (2001) identifies “discourse-oriented historical linguistics” as having originated with Givón’s (1979) identification of the “historical progression: discourse > syntax > morphology > morphophonemics > zero” (Brinton 2001: 147). In this context, DMs are particularly interesting for investigations of language change (Brinton 2001: 152). DMs are also important for “diachronically oriented discourse analysis” (Brinton 2001: 147).

Although I have argued elsewhere that conversation is the matrix of discourse (Brody 1986, 2006), here I follow the general Americanist tradition of focusing on narrative texts in the investigation of indigenous languages (Darnell 1990). In particular, I follow Kwachka’s (1992: 70) observation that indigenous people “have been able to transfer and permute a very important cultural pattern at the discourse level – the tradition of narrative”. Narrative tradition is a powerful source of memory in languages with no written tradition, such as Tojol-ab’al. Hence I have chosen to trace DMs as they are used in Tojol-ab’al narrative over time.

I have identified five major functions for Tojol-ab’al DMs based on a modification of Fraser’s (2006) classification: elaborative, contrastive, temporal, inferential and “change of topic or speaker, introduction or termination of a discourse unit” (Brody 2010: 20). The latter function was added to account for the operation of DMs, both indigenous and borrowed, in Tojol-ab’al; this illustrates the fact that some DMs borrowed from Spanish have taken on the function of indigenous DMs. The data below shows the operation of all of these functions across the three generations, as well as the frequency of occurrence of both borrowed and indigenous DMs. We will see that the use of both borrowed and indigenous DMs remains for the most part the same across the three generations investigated, demonstrating a stability of Tojol-ab’al discourse structure through time.

## 2. The data

The data consists of transcripts of audio recordings of speech of eight individuals from three generations of speakers of Tojol-ab’al. The oldest generation of speakers (now deceased) is represented by the speech of two sisters, one of whom recently died at the age of 117; they were monolingual in Tojol-ab’al, and I recorded them in 1980. The middle generation is comprised of three speakers now in their 60s, recorded in the early 1990s; there are two male speakers who are sons of the sisters, and the female speaker is married to one of the sons. The men received schooling

in Spanish as adults, but the woman never attended school. The youngest speakers all attended school (in Spanish) starting in elementary grades. This group includes two women and one man in their late teens and early 20s, recorded in 2008 by the daughter of one of the middle generation males from among her network of friends.<sup>1</sup>

Although no associated sociolinguistic survey was made, each of the individuals represented identifies as a speaker of Tojol-ab'al. All of the speech recorded is from the genre of conversational narrative, on topics selected by the speakers. Although the texts are interactive, the word counts represent only the speech of the dominant speaker in these conversational narratives.

### 3. Methods

The texts were recorded, transcribed and translated with the assistance of experienced bilingual consultants. Out of a larger corpus, I selected narrative texts that were as comparable as possible in subject matter and would yield approximately equivalent word counts for each generation. Since the recordings, transcripts and translations were made over a period of time, I reviewed all of the texts to ensure that a consistent transcription system was used throughout. I then made a count of the total number of words in each text and enumerated the indigenous DMs (see Figure 1), those borrowed from Spanish (see Figure 2; compare with the list in Brody 2010 from a different corpus) and all other words borrowed from Spanish. Names of people, places and saints were not counted as borrowed. Percentages of indigenous DMs, borrowed DMs and other borrowed words were calculated for all of the texts.

Eldest generation: *ti* “then” (37), *yajni* “when” (12), *jach* “thus” (7), *cho* “and” (5), *jastal* “how” (2), *jawa* “that” (2), *ja`yuj* “because” (1) [Total 66; 3.9%]

Middle generation: *ti* “then” (23), *cho* “and” (13), *yajni* “when” (9), *jaw* “that” (8), *jastal* “how” (2), *jach* “thus” (2), *sok* “and” (1) [Total 58; 6.2%]

Youngest generation: *ti* “then” (6), *cho* “and” (4), *sok* “and” (4), *yujni* “because” (2), *jach* “thus” (2), *jawa* “thus” (2), *jasa* “thus” (1) [Total 21; 1.7%]

Figure 1. Inventory of Tojol-ab'al DMs by generation

1. Caroline Covington participated in this part of the project.

Eldest generation: *pues* Sp. *pues* “then” (27), *como* Sp. *como* “since” (26), *porke* Sp. *porque* “because” (13), *i* Sp. *y* “and” (13), *entonse/antonse* Sp. *entonces* “then” (11), *pero* Sp. *pero* “then” (7), *kwando* Sp. *cuando* “when” (5), *ke* Sp. *que* “what” (3), *asta ora* Sp. *hasta ahora* “until now” (4), *este* Sp. *este* “um” (2), *no* Sp. *no* “no” (1), *si* Sp. *si* “if” (1), *bweno* Sp. *bueno* “well” (1), *asi* Sp. *así* “thus” (1), *miyentras* Sp. *mientras* “meanwhile” (1), *por eso* Sp. *por eso* “for that reason” (1), *eso* Sp. *eso* “so” (1) [Total 118; 7.1%]

Middle generation: *entonse/antonse* Sp. *entonces* “then” (18), *i* Sp. *y* “and” (13), *pwes* Sp. *pues* “well” (9), *pero* Sp. *pero* “but” (13), *porke* Sp. *porque* “because” (3), *komo* Sp. *como* “since” (3), *iday* Sp. *y de ahí* “and then” (3), *este* Sp. *este* “um” (2), *bweno* Sp. *bueno* “well” (2), *eske* Sp. *es que* “it’s that” (2), *sike* Sp. *si que* “if that” (1), *tambiyen* Sp. *también* “also” (1), *ke* Sp. *que* “that” (1) [Total 70; 7.5%]

Youngest generation: *i* Sp. *y* “and” (39), *pero* Sp. *pero* “but” (14), *este* Sp. *este* “um” (7), *porke* Sp. *porque* “because” (6), *ke* Sp. *que* “that” (5), *pwes* Sp. *pues* “well” (4), *tambiyen* Sp. *también* “also” (3), *o* Sp. *o* “or” (3), *poreso* Sp. *por eso* “for that” (3), *parake* Sp. *para que* “so that” (1), *komo* Sp. *como* “since” (1), *bweno* Sp. *bueno* “well” (1), *por ejemplo* Sp. *por ejemplo* “for example” (1), *este* Sp. *este* “um” (1) [Total 89; 7.5%]

Figure 2. Inventory of borrowed DMs by generation

Word count is somewhat complicated for this language by the fact that Tojol-ab'al has polysynthetic tendencies in its verbs and otherwise complex morphology with numerous clitic particles. A consistent word division was maintained across the transcripts examined here to facilitate comparability across speakers. When a Spanish word was incorporated with Tojol-ab'al morphology (e.g., *s-parte* “su parte”, “his part”) it was still counted as a borrowed word. Each token of use of a borrowed word was counted; this is especially important because of the high degree of repetition that characterizes Tojol-ab'al discourse (Brody 1986); repetition, I argue, can itself function as a DM in Tojol-ab'al (Brody 2010, 2015). To determine what counted as a Spanish DM, Poblete's (1998) inventory was used; Tojol-ab'al DMs were determined by function (Brody 1987, 1989, 1995, 2010). I calculated percentages of borrowings of DMs, of indigenous DMs, of other borrowed words and the total combined borrowings for the texts of each generation; because the sample was so small (only 8 speakers total), no further statistical analysis of the findings was carried out. I also examined the texts from each generation of speakers for examples of the five functions of borrowed DMs listed above.

## 4. Findings

### 4.1 The eldest generation

There are two texts from the eldest generation, one from each sister. The shorter text (229 words) recounts a procedure of tasks undertaken after the birth of a

baby in traditional times. The longer text (1432 words) is a reminiscence of the women's early life, and it contains a good deal of reported discourse, including a sequence of advice given to a new bride. Reported discourse is a common feature of Tojol-ab'al narratives and is included in the word counts. Combining these two texts, there are 118 borrowed DMs (7.1% of total words) and 120 other tokens of words borrowed from Spanish (7.2% of total words). The percentage of tokens of borrowed DMs thus closely equaled the percentage of tokens of other words borrowed from Spanish. Together, the borrowings account for 14.3% of the total words.

The procedural text uses the indigenous DM *ti* "then" most frequently (14 times) and the borrowed DM *antonse/entonse* – Sp. *entonces* "then" – (7 times) with the greatest frequency as a factor of its content; it describes a series of activities, each occurring in a temporal sequence, as shown in example (1) below. All borrowings are underlined in the data examples, and all DMs are italicized. We see in (1), line 5, the paired use of indigenous and borrowed DMs.<sup>2</sup>

- (1) 3 *ti* wa x-b'oy-x-i-0 b'a y-oj s-nalan  
 then INC INC-bury-MID-IVM-3A LOC 3E-interior 3E-half  
 ja k'ak'=i  
 DET fire=TERM  
 "then the center of the fire gets dug up"
- 4 *ti* wa x-muk-x-i-0=a  
 then INC INC-bury-MID-IVM-3A=TERM  
 "then it gets buried"
- 5 *entonse* *ti* wa x-ch'ak k'ak'=a ja s-muxuk'=i  
 then then INC INC-finish burn=TERM DET E-placenta=TERM3  
 "then then the placenta burns up" (Babies)<sup>3</sup>

The reminiscence text uses *pwes* – Sp. *pues* "then" – most frequently (27 times), usually in the function of initiating of a new topic or a new speaker. Both uses are shown in example (2) below. CH introduces a new topic in 385 and then begins to quote another speaker {AB} in 386. Both lines also shows pairing of

2. Abbreviations used are as follows: 1A, 2A, 3A – first-, second- and third-person absolutive; 1E, 2E, 3E – first-, second- and third-person ergative; CL – cleft marker; COM – completive; DET – determiner; INC – incomplete; IRR – irrealis; IVM – intransitive verb marker; LOC – locative; MID – middle voice; NEG – negative; PL – plural; PN – pronoun; PRO – progressive; REP – reportative; SBJ – subjunctive; TERM – terminal; TVM transitive verb marker.

3. Names in parentheses after examples indicate the text from which the example is taken.

indigenous DM *yahni* “when” (line 385) and *tixa* “then” (line 386) and the borrowed DM *pwes* “pues”.

- (2) 385 CH: *pwes* *yajni* 0-waj-i-on=i  
 well when COM-leave-IVM-1A=TERM  
 “well when I left”
- 386 “*pwes* *ti=xa* a-wil a-b’aj=a”  
 well then=now 2E-see 2E-SELF=TERM  
 “well then you will see” {AB} (Chica)

This text also provides an example of the use of the inferential function with the borrowed DM *porke* – Sp. *porque* “because” – as seen in example (3).

- (3) 114 CH: oj k-a`- y-i`-0 s-wa`el  
 IRR 1E-give- -3E-take=3A 3E-food  
 “I will give him food”
- 115 *porke* j-kijts’in=i  
 because 1E-younger.sibling=TERM  
 “because he is my little brother” (Chica)

Finally, the contrastive function with *pero* – Sp. *pero* “but” – and the temporal functions with *ora* – Sp. *ahora* “now” – are illustrated in the sequence in example (4) from the reminiscence text. Line 211 shows pairing of the indigenous DMs *yaj ora* “when” and *pero* – Sp. *pero* “but”.

- (4) 210 CH: *j-parte* k-e`n-tik-on j-nupan-uk-0 ja  
 1E-parte 1E-PN-PL-1A 1E-marry-SBJ-3A DET  
 ala k-ijts’in=i  
 dear 1E-younger.sibling=TERM  
 “on our part we married my little brother”
- 211 *pero* ja *yaj ora* el-i-on b’a s-parte  
 but DET when when leave-IVM-1A LOC 3E-parte  
 ja j-me`xep=i  
 DET 1E-grandmother=TERM  
 “but when when I left the care of my grandmother”
- 212 ch’in-0=to 0-k-a`-0 kan=i  
 small-3a=still COM-1E-give-3A stay=TERM  
 “I left when he was still small” (Chica)

Thus all five functions of DMs are demonstrated for the eldest generation.



## 4.2 The middle generation

The middle generation is represented by three texts, two of which are from sons of the two elder generation sisters; the third is from the wife of one of the sons. One text is a personal reminiscence (183 words), and two are stories, all related in conversation. One story is from one of the men (407 words), the other from the wife of the man who told the personal reminiscence (340 words), for a combined total of 932 words. These individuals are all bilingual in Tojol-ab'al and Spanish. The two men are fully bilingual, but the woman's Spanish is weaker than the men's, and her speech was anomalously conservative: she had only 9 borrowed DMs in 340 words, comprising just 2.6% of her total words; this is an even lower percentage than that used by the older generation. Her weakness in Spanish probably accounts for this low number of borrowed DMs. However, taken as a group, the middle generation showed a very similar rate of overall borrowed words – 14% – to the older generation's 13.9%, with borrowed DMs (72 or 7.7%) and other borrowed words (59 or 6.3%) occurring also at a closely equivalent rate.

Although all of the texts from the middle generation relate sequential actions, only one of them has a high rate of borrowed DMs indicating temporal action: *i* – Sp. *y* “and” (11 instances) – and *entonse* – Sp. *entonces* “then” (18 instances). A series of examples is seen in the sequence in (5).

- (5) 3 *i* ja y-alajil ja s-chenek'=i  
 and DET 3E-field DET 3E-beans=TERM  
 jel x-ch'ak-0 chan  
 very INC-finish-3A animal  
 “and her field of beans was really being done in by animals”
- 4 *i* mi s-na`-a-0 jasu`a ja  
 and NEG 3E-know-TVM-3A what DET  
 wan y-a`-jel- -y-i`-0  
 INC 3E-give-PRO-3A- -3E-take-3A  
 “and she didn't know what was doing it in”
- 5 *i* sk'ulan jun manya  
 and 3E-make-3A one artifice  
 “and she made an artifice” (Birjin)

In (6) we see an example from the woman's story of a borrowed DM indicating change of topic. This example also shows the pairing of borrowed (*pwes* – Sp. *pues* “then”) and indigenous (*ti* “then”) DMs.

- (6) 254 J: *pwes* *ti* kayan 0-ek'=a jaw=i  
 well then float COM-pass=TERM that-TERM  
 “then then that one was floating along” (2x2 crazy)

Another example from the woman's story, in (7) below, shows the elaborative function with *eske* – Sp. *es que* “it's that”.

- (7) 365 J: *eske*    *jel=b'i*    *mal*    *juntiro*    *ja*    *wa*    *x-k'e-0*    *lokoil*  
          it's.that    very=REP    bad    much    DET    INC    INC-raise-3A    crazy  
          *jastal*    *jaw=i*    *rabiyo=i*  
          like    that=TERM    rabid=TERM  
          “It's that supposedly it is very bad indeed when we get crazy like  
          that, rabid” (2x2 crazy)

The following example in (8), from one of the men's stories, shows both contrastive function (with *pe* – Sp. *pero* “but”) in line 4 and inferential function (with *porke* – Sp. *porque* “because”) in line 3.

- (8) 3 L: *porke*    *kechan*    *j-tuch'il*  
          because    only    1E-alone  
          “because I was alone”  
      4 *pe*    *ja=xa*    *mas*    *tza'an*    *0-el-i-on=a*  
          but    DET=now    more    later    3E-leave-IVM-1A=TERM  
          “but as for later on I left” (Lendtxt)

Thus we have all five functions of DMs occurring in the speech of the middle generation as well.

### 4.3 The younger generation

There are three texts from the younger generation. All are accounts of personal life and life in the community. One is from a young man (441 words) and two are from young women (373 and 371 words), for a total of 1185 words. As was the case for the two elder generations, the topics the young people discussed were chosen by the speakers. All of the younger speakers have attended school and are fluent in Spanish; however, they identify as speakers of Tojol-ab'al.

Looking at only the borrowed DMs, the percentage in this age group is 7.6%, which is very similar to that of the two preceding generations. However, one of the most striking aspects of the speech of the youngest generation is the higher rate of borrowing in general. Counting both DMs and other words, 24.6% of the total words in the speech of the younger participants were borrowed from Spanish. Part of this is due to subject matter, as the young people discuss their educational and recreational activities, for example, for which Tojol-ab'al has no words, such as *universidad* “university”, *fútbol* “soccer”, etc. Part of the higher Spanish use was also due to the tendency to repeat phrases in Spanish, e.g., *lo que sea* “whatever”, or *de vez en cuando* “from time to time”. As mentioned above, repetition can also

function as a DM of changing topic in Tojol-ab'al (Brody 2010, 2015). Another minor factor in the greater use of Spanish by the younger generation was the spoken IRB consent protocol used, where the Spanish term *grabar* 'record' and the date in Spanish occurred and were repeated by the participants. Yet these speakers also sometimes used Spanish words where there are Tojol-ab'al equivalents, for example, for numbers.

Another notable difference in the speech of the younger generation is their lower use of indigenous DMs: 1.7% as compared with 6.2% for the middle generation and 3.9% for the eldest generation.

The elaborative and temporal functions of the borrowed DMs *porke* – Sp. *porque* 'because' – and *i* – Sp. *y* 'and' – are shown in the sequence in example (9).

- (9) 36 V: *jel* *x-k'ulan=ni* *j-gusto* *oj* *ajyik-on* *sok*  
 very INC-like-EMP 1E-pleasure IRR be-1A with  
*ja* *k-amigos*  
 DET 1E-friends  
 'I really like to be with my friends'

- 37 *porke* *wa* *x-k-a'-0* *tsenok-e'*  
 because INC INC-1E-give-3E laugh-PL  
 'because I make them laugh'

- 38 *i* *wa* *x-j-k'ul-a-tikon* *jitsan* *kosa*  
 and INC INC-1E-do-TVM-3A.PL many thing  
 'and we do a lot of things'

(Forma de Vivir)

The cluster of borrowed discourse markers in example (10) (line 47) illustrates change of topic.

- (10) 46 V: *i* *jel* *gusto* *waj-0* *ek'* *k-uj-tikon*  
 and very pleasure go-3A pass 1A-agency-1PL  
 'and we have a good time'

- 47 *i* *tambiyen* *este* *ja* *ja* *b'a* *chonab'*  
 and also um DET DET LOC Comitán  
 'and also um in in Comitán'

(Forma de Vivir)

The elaborative function is shown in example (11) with the use of *porke* – Sp. *porque* 'because'.

- (11) 43 M: *porke* *ja* *il=i* *jel* *x-k-ab'-jel*  
 because DET here=TERM very INC-1E-feel-PRO  
*wan-on* *yijel*  
 PRO-1A suffer  
 'because here I feel I am suffering'

(a'tel)

An example of the contrastive function with *pero* – Sp. *pero* “but” – is seen below in (12) (line 11).

(12) 9 V: wan-i-0 j-k’uman-i-tikon ja Tojol-ab’al  
 PRO-IVM-3A 1E-speak-IVM-1A.PL DET Tojol-ab’al  
 “we do speak Tojol-ab’al”

10 ama seya mi ja’-uk-0 ja mero=i  
 NEG if NEG CL-SBJ-3A DET real=TERM  
 “even though it is not the real thing”

11 pero wa la-k’umani-tikon  
 but INC INC1A-speak-1A.PL  
 “but we speak it”

(Forma de Vivir)

I include this example not only to show the borrowing of *pero* as a contrastive DM (*seya* – Sp. *sea* “even though” – is also borrowed in line 10, although not as a DM), but also because it expresses a strongly held attitude in the Tojol-ab’al language community: young people no longer speak Tojol-ab’al well. That this sentiment is voiced in the speech of a young man speaking in Tojol-ab’al to an age-mate shows how deeply this attitude is engrained (see Brody 2010 for a discussion of borrowed DMs in the context of local language ideologies). Despite the lower rate of DMs in general, the youngest generation also uses all five functions of the DMs.

## 5. Discussion

All of the borrowed DMs in the corpus follow the typical pattern of borrowing as documented in Brody (2010): they are utterance-initial, they always precede the indigenous DM and they are adapted to Tojol-ab’al phonology (e.g., *antonse* – Sp. *entonces* “because”). Furthermore, all of the discourse functions for both indigenous and borrowed DMs were found in each generation’s speech (not all possibilities are shown in the examples, but see Figures 1 and 2).

The youngest generation shows a marked change from the two previous generations in its general pattern of borrowing. Borrowing from Spanish was much greater altogether in the speech of the youngest generation, with all words from Spanish accounting for 24.6% of their speech. I am assuming that this use of Spanish is a result of borrowing rather than code-switching due to the powerfully puristic linguistic ideology of Tojol-ab’al people to speak only one language at a time (Brody 2010). However, it remains to be determined if that language ideology has altered with the youngest generation, making it possible that their use of Spanish is code-switching. This is a question for future research. Nevertheless, their use of

borrowed DMs was consistent with that of the previous generations, with 7.5% of the words representing borrowed DMs.

With regard to indigenous DMs, the inventory for all three generations is somewhat different for each generation, but there is a great deal of overlap. In the corpus considered here, the youngest generation used indigenous DMs as only 1.7% of their total words, as compared to 6.2% for the middle generation and 3.9% for the older generation (see Figure 1). The expansion in use of indigenous DMs in the middle generation over that of the eldest generation seems to run counter to the expectation of loss through time. Although the younger generation's percentage is somewhat lower than that of the eldest generation, the important fact is that we do not see a complete loss through time of indigenous DMs.

The inventory of borrowed DMs (see Figure 2) for the middle generation is slightly smaller (14 items) than is that of the elder generation (17 items). The younger generation used 15 different borrowed DMs. For the youngest speakers, the most commonly used by far of the borrowed DMs is *y*, mirroring its high frequency as a DM in Spanish (Poblete 1998).

The general tendency of borrowed DMs to occur alongside indigenous DMs (Aikenvald 2002; Heine & Kuteva 2005; Muysken 2000) is amply demonstrated in Tojol-ab'al (Brody 2010) and seen in examples (1) (line 5), (2) (lines 385 and 386), (4) (line 211) and (6) (line 354) above and in example (13) below.

The youngest generation also uses sequences of indigenous DMs as do the elder generations, as shown in (13), where the borrowed DMs *i* – Sp. *y* “and” – and *poreso* – Sp. *por eso* “for this reason” – co-occur with the indigenous DM *ja'yuj* “because”.

- (13) 48 M: *i poreso ja'-0 y-uj k-i'-onej*  
 and for.this.reason CL-3A 3E-agency 1E-take-PRO-3A  
*esas ganas*  
 these desires  
 “and for this reason is why I have these desires” (a'tel)

The younger generation also pairs up borrowed DMs, as seen in (14) below.

- (14) 18 V: *i tambiyen jel j-k'ulan-0 j-gusto*  
 and also very 1E-do-3A 1E-pleasure  
 “and also I like it a lot” (Forma de Vivir)

## 6. Conclusions

The most striking finding of this study is that, despite other changes, the rate of borrowed DMs has remained constant at slightly more than 7% over the three

generations. It is also clear that DMs continue to exhibit the same functions across the three generations. Despite the lowered numbers of indigenous DMs in the speech of the youngest generation, they still exhibit all of the functions of borrowed and indigenous discourse markers as do the speakers in the elder two generations, as well as the use of pairing of borrowed and indigenous DMs. Thus, I claim that the basic discourse structure of Tojol-ab'al, as indicated by the use of DMs, has remained essentially the same across the span of three generations.

It might seem on the surface that the language attitude articulated by the young man quoted above in example (12) – that the Tojol-ab'al he speaks “is not the real thing” – is borne out to a degree, in that nearly 25% of the total words in the younger generation's speech are in Spanish – a much higher rate than that of the two older generations – and only 1.7% of the total words are indigenous DMs – much lower than the older two generations. However, as noted above, the increase in Spanish is in large part a result of the subject matter that the younger people discuss, topics for which there are no indigenous words in Tojol-ab'al. This is a variable that could not be controlled, due to the naturalistic nature of the data and the fact that speakers chose their own topics of conversation. While the younger generation uses a lower percentage of indigenous DMs, they exhibit a similar inventory of indigenous DMs to the older generations. It may turn out that as they age, the speech of the younger generation might come to more closely resemble that of their elders; alternatively, they may continue to use even more Spanish. There are other differences, beyond the scope of this paper, which also indicate that the speech of the youth has changed from that of the preceding generations. Yet my analysis here shows that Tojol-ab'al discourse organization, as structured by both indigenous and borrowed DMs, remains markedly the same across the three generations. This finding contradicts the conclusions of Clyne (2003) and Matras (1998), who found that borrowing of DMs led to changes in discourse structuring over time.

Of course three generations is a very short time span in the context of 450 years of contact. It could be argued that changes in Tojol-ab'al discourse due to borrowed DMs have already taken place long ago. Yet, as mentioned above, there are many features of Tojol-ab'al discourse that are quite distinct from Spanish discourse, indicating that any influence must be minor. Furthermore the three generations represented here lived through a time period of turbulent change for the Tojol-ab'al community. Much of the Tojol-ab'al speaking region was until recently quite remote from the rest of Mexico and even from the rest of the state of Chiapas due to the lack of roads (Collier & Quintanillo 2005). Changes that had an impact on the middle generation include the incursion of Protestant missionaries into the region and resulting expulsion of converts from their native villages; refugees often fled to Spanish-dominant areas (Brody 2006; Collier & Quintanillo 2005).

Relevant to the youngest generation, it was only after recent events prompted by the Zapatista rebellion of 1994 that more of the region become accessible by newly paved roads, bringing in more (Spanish language oriented) schools as well as other changes from the Spanish-speaking world (Brody 2001, 2010).

I suggest also that the continuity found in Tojol-ab'al DMs is perhaps due in part to the continued practice of pairing of borrowed and indigenous DMs across all three generations. The Tojol-ab'al preference for repetition (Brody 1986, 2010) serves to reinforce the use of paired borrowed and indigenous DMs. Although this pilot study investigates a very small number of speakers – only eight in total – it is noteworthy that we are able to examine the speech of three generations of indigenous Tojol-ab'al speakers to compare their language use and see signs of stability in the face of long-term contact with and increased borrowing from Spanish.

## References

- Aikenvald, Alexandra Y. 2002. *Language Contact in Amazonia*. Oxford: Oxford University Press.
- Brinton, Laurel J. 2001. "Historical Discourse Analysis". *The Handbook of Discourse Analysis* ed. by Deborah Schiffrin, Deborah Tannen, & Heidi E. Hamilton, 138–160. Malden, Mass.: Blackwell.
- Brody, Jill. 1986. "Repetition as a Rhetorical and Conversational Device in Tojol-ab'al (Mayan)". *International Journal of American Linguistics* 53.255–274. doi:10.1086/466022
- Brody, Jill. 1987. "Particles Borrowed from Spanish as Discourse Markers in Mayan Languages". *Anthropological Linguistics* 29.175–184.
- Brody, Jill. 1989. "Discourse Markers in Tojol-ab'al Mayan". *Chicago Linguistic Society Parasession on Language in Context* ed. by Bradley Music, Randolph Graczyk & Caroline R. Wiltshire, 15–29. Chicago: Chicago Linguistic Society.
- Brody, Jill. 1995. "Lending the 'Unborrowable': Spanish Discourse Markers in Indigenous American Languages". *Spanish in Four Continents: Studies in Language Contact and Bilingualism* ed. by Carmen Silva-Corvalán, 132–147. Washington: Georgetown University Press.
- Brody, Jill. 1998. "On Hispanisms in Elicitation". *Convergencia e individualidad: Las lenguas mayas entre hispanización e indigenismo* (Colección Americana No. 7, Universität Bremen) ed. by Andreas Koechert & Thomas Stoltz, 61–84. Hanover & Guatemala City: Verlag für Ethnologie.
- Brody, Jill. 2001. "From Conquistadores to Zapatistas: Language Contact and Change in Tojol-ab'al". *Southwest Journal of Linguistics* 20.1–17.
- Brody, Mary Jill. 2006. "Responsibility in Tojol-ab'al Gossip: Indirect Speech, Modal Orientation, and Metalinguistic Terms as Used to Construct Self and Other in a Moral Landscape". *Ketzalcalli* 2.2–21.
- Brody, Mary Jill. 2010. "'Sticky' Discourse Markers in Language Contact Between Unrelated Languages: Tojol-ab'al (Mayan) and Spanish". *A New Look at Language Contact in Amerindian Languages* ed. by Claudine Chamoreau, Zarina Estrada Fernández & Yolanda Lastra, 9–36. Munich: Lincom Europa.

- Brody, Mary Jill. 2015. "Exact Repetition in Tojol-ab'al". Paper presented at the Workshop on Exact Repetition, Leipzig, March 2015.
- Cepeda, Gladys & María T. Poblete. 1996. "Marcadores conversacionales: Función pragmática y expresiva". *Estudios filológicos* 31.105–117.
- Clyne, Michael G. 2003. *Dynamics of Language Contact: English and Immigrant Languages*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511606526
- Collier, George & Elizabeth R. Quaratiello. 2005. *Basta! Land and the Zapatista Rebellion in Chiapas*, 3rd ed. Oakland: Food First Books.
- Darnell, Regna. 1990. "Franz Boas, Edward Sapir, and the Americanist Text Tradition". *Historiographia Linguistica* 17:129–144. doi:10.1075/hl.17.1-2.11dar
- Fraser, Bruce. 2006. "Towards a Theory of Discourse Markers". *Approaches to Discourse Particles* ed. by Kerstin Fischer, 189–204. Amsterdam: Elsevier.
- Givón, Talmy. 1979. "From Discourse to Syntax: Grammar as a Processing Strategy". *Syntax and Semantics*, vol. 12: *Discourse and Syntax* ed. by Talmy Givón, 81–112. New York: Academic Press.
- Grimes, Joseph E. 1972. "Outlines and Overlays". *Language* 48.513–524. doi:10.2307/412030
- Haugen, Einar. 1950. "The Analysis of Linguistic Borrowing". *Language* 26.210–231. doi:10.2307/410058
- Heine, Bernd & Tania Kuteva. 2005. *Language Contact and Grammatical Change*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511614132
- Kwachka, Patricia. 1992. "Discourse Structures, Cultural Stability, and Language Shift". *International Journal of the Sociology of Language* 93:67–73. doi:10.1515/ijsl.1992.93.67
- Lenkersdorf, Gudrun. 1986. "Contribuciones a la historia colonial de los tojolabales". *Los legítimos hombres: Aproximación antropológica al grupo tojolabal*, vol. 4, ed. by Mario H. Ruz, 13–102. Mexico City: Centro de Estudios Mayas.
- Matras, Yaron. 1998. "Utterance Modifiers and Universals of Grammatical Borrowing". *Linguistics* 36.281–331. doi:10.1515/ling.1998.36.2.281
- Matras, Yaron. 2006. "The Borrowability of Structural Categories". *Grammatical Borrowing in Cross-linguistic Perspective* ed. by Yaron Matras & Jeanette Sakel, 31–74. Berlin: Mouton de Gruyter.
- Muysken, Pieter. 2000. *Bilingual Speech: A Typology of Code-Mixing*. Cambridge: Cambridge University Press.
- Poblete, María T. 1998. "Los marcadores discursivo-conversacionales de más alta frecuencia en el español de Valdivia (Chile)". *Estudios filológicos* 33.93–103.
- Schiffrin, Deborah. 1987. *Discourse Markers*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511611841
- Stolz, Christel & Thomas Stolz. 1996. "Funktionswortentlehnung in Mesoamerika. Spanisch-amerindischer Sprachkontakt Hispanoindiana II." *Sprachtypologie und Universalien-Forschung* 49.86–123.
- Thomason, Sarah G. & Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Berkeley: University of California Press.
- Torres, Lourdes. 2006. "Bilingual Discourse Markers in Indigenous Languages". *International Journal of Bilingual Education and Bilingualism* 9.615–624. doi:10.2167/beb383.o





## CHAPTER 6

### *Hasta perder la última gota de mi sangre*

#### Variation in commissive speech act behavior in Colonial Louisiana Spanish

Jeremy King

Louisiana State University

In spite of the recent burgeoning of the field of Spanish socio-pragmatics, the category of commissive speech acts has been all but ignored in scholarship, particularly in contact varieties of the language. Colonial Louisiana presents an ideal context for this type of work. The analysis herein involves a corpus of 100 business letters between Spanish government and military officials in 18th century Louisiana. All commissive speech acts were identified and coded according to a modified version of the taxonomy presented in Bilbow (2002). Results of the study reveal that the level of power of the speaker (= letter writer) proved to be the most significant factor in the formulation of commissives. The level of relative power of a speaker appeared to guide not only the general type of commissives employed (e.g., promises vs. offers) but also the specific morphological expression of the speech act (e.g., future vs. present tense).

#### 1. Introduction

Ever since the advent of speech act theory by language philosophers such as Austin (1962) and Searle (1976), the notion that we DO things when we produce language has remained at the forefront of the field of linguistic pragmatics. In spite of the caution issued by some researchers (Dippold 2007, for instance) that the study of individual speech acts is insufficient to capture a clear picture of a speaker's pragmatic intent, speech act analysis has remained one of the most popular aims of pragmatics research due to the fact that these elements are said to "constitute the 'basic' or 'minimal' units of linguistic communication" (Márquez Reiter & Placencia 2005: 16). Although some speech act types have received significant attention in the pragmatics literature, commissive speech acts have been sparsely

examined in empirical work, particularly from an historical perspective. The current study seeks to expand the literature on commissives in relation to the Spanish language in particular, as well as to document the usage of this linguistic feature in colonial Louisiana in the second half of the 18th century.

Commissives are defined as those speech acts which commit the speaker to a certain course of action (Searle 1976: 7). Upon uttering (or, in more general terms, producing) a commissive speech act, a speaker commits him/herself to a future action. It is generally accepted among language theorists that the uttering of a commissive binds the speaker to the course of action described and creates expectations in the hearer that said action will be carried out. These acts can be expressed transparently through the use of performative verbs such as *promise* and *vow*, or commitment can be offered through a wide range of linguistic means, including minimal responses to questions such as “Sure” and “OK” (Bilbow 2002: 296). One of the most salient, and most frequently theorized, aspects of commissives is that of the so-called *sincerity condition* involved in their production; with commissives, this issue is particularly important, given that the focus in most settings of everyday communication is placed on the integrity and candor of the speaker when committing to an action. Csordas (1997: 325) asserts that this focus is drawn from the genuine human need to ‘contract’ social bonds and act sincerely in our undertakings.

In the current study, I examine the expression of commissive intent in a corpus of business letters taken from the Louisiana territory in the latter half of the 18th century, when this portion of North America was under Spanish control. I consider the ways in which commissives were formulated in this period as compared to extant data on both modern varieties of Spanish as well as in modern studies of business communication in general. The specific focus of this chapter is the ways in which the variable of institutional power affects the communication of commissive force in a business setting.

## 2. Commissive speech acts and the Spanish language

Unlike many different types of speech acts, such as directives and expressives, commissives have been a relatively unpopular focus of linguistic research, and even more so when it comes to empirical studies. One potential reason for the dearth of research on this category of speech acts may be found in their very nature. On the one hand, commissives may not be as easily accessible via spontaneous discourse as other speech acts; because of this, many studies that focus on this speech act type use data gained via a variety of elicitation methods, as will be seen in this section. On the other hand, Leech (1983) and Haverkate (1994) maintain that

commissives are inherently polite speech acts, given that their successful fulfillment benefits the hearer in an interaction. If this is the case, the study of commissives might seem less alluring to linguists than that of other acts, as they are less threatening and thus less controversial. Putting aside that the notion of ‘intrinsically’ polite speech acts has been highly debated in the pragmatics literature (cf. Brown and Levinson 1987; Fraser and Nolen 1981; Watts 2003; etc.), a number of authors have cautioned against labeling commissives in this way. Chodorowska-Pilch (2002), for instance, argues that offers, one specific instantiation of commissives, are formulated with mitigation strategies in some contexts of language use, which lends support to their classification as potentially face-threatening acts.

Much of the extant work involving commissives takes the form of language philosophy à la Austin/Searle (cf. Hancher 1979; Haugh 2008; Holtgraves 2005; Huang 2010; Johansson 2003; Pérez Hernández 2001; Pérez Hernández & Ruiz de Mendoza 2002; Verschueren 1983; Sbisà 2001). In addition, many authors have examined specific facets of this speech act type from the perspective of a wide variety of fields, including law (Danet 1980; de Jong & Werner 1998; Tiersma 1986, 1993; Trosborg 1995; Visconti 2009), religion (Bruder 1998; Csordas 1997; Desilva 1994; Larson 1978; Wisse 2000), literary studies (Bernaerts 2010; Boix Jovaní 2012; Fletcher 2003; McCallam 2003; Person 2009; Polansky 1988; Rudanko 2004; Sánchez 1980), sociology (Brewer 2003), business language (Bilbow 2002), archaeology (Byers 1999), technology (Carvalho & Cohen 2005; Sitter & Stein 1992); politics (Knops 2006; McDonald 1969) and even medicine (Natsopoulos et al., 1993). Notwithstanding the significant quantity of research involving commissives, relatively little empirical work has been carried out on this class of speech acts within linguistics. A handful of studies have focused on commissives in language acquisition contexts (Aston 1988; Bernicot & Laval 1996; Olson 2007; Félix-Brasdefer 2003) and language history and development (Myhill 1996), and several descriptions of this speech act type in non-Western languages have been offered (Agyekum 2004; Liu 2011; Rosaldo 1982), but these studies are few and far between.

Work on commissives in the Spanish language has likewise been sparse. In their discussion of this speech act category, Márquez Reiter & Placencia (2005: 74–75) note that, as of their time of writing, only three pragmatics-related studies had been carried out on commissives in Spanish. To date, a number of additional publications on the topic have appeared; these studies have used vastly different corpora as their points of departure and have focused on different aspects of commissives. In an early study, Polansky (1988) analyzes the functions of commissives in a literary corpus, specifically the well-known 19th century dramatic play *El desengaño en un sueño*. The author holds that the structure of the play is balanced on two ultimately felicitous commissive illocutionary

acts. She further argues that Lisardo, the protagonist, invokes strategic challenges via the use of commissives, which ultimately serve to expedite his wishes being fulfilled (1988: 7). This finding is strikingly similar to the conclusions drawn by McCallam (2003) in a linguistic analysis of the 18th century French work *Les liaisons dangereuses*. Thus, it can be concluded that, for at least the past several centuries, speakers of Spanish (and perhaps European languages in general) have recognized that commissive speech acts can function as tools of influence and manipulation and have exploited this.

Hardin (2001) investigates the use of distinct types of speech acts in Spanish language television advertising. She notes that, in her corpus of speech act data, commissives make up fewer than 20% of the total tokens counted. This number seems somewhat low, considering that the focus of Hardin's study is persuasive discourse within the advertisements she analyzes. Importantly, despite the frequent use of overt performative verbs to express commissive force (*ofrecer* "to offer", *regalar* "to gift", *garantizar* "to guarantee", etc.), the author points out that the commissives used in advertising are generally qualified in some way and often require action on the part of the viewer to 'activate' the action to which the advertiser is committing. Bilbow (2002: 293) refers to utterances such as these as "conditional commissive speech acts", in which the speaker seeks to share, rather than assume, responsibility for an action. Trosborg (1995: 36) asserts that the use of explicit statements is desirable when producing commissives, as indirectness or hedging would only serve to weaken the obligation inherent in the speech act and potentially render the act less binding. Thus, attenuated commissives of the type noted by Hardin are dispreferred and marked forms.

Chodorowska-Pilch (2002) discusses offers in Peninsular Spanish vis-à-vis politeness theory. In her data collected from travel agencies in Spain, the author notes several salient points about offers in this variety of Spanish: first, offers are nearly always expressed indirectly (i.e., typically not with performative verbs such as *ofrecer* "to offer" or *proponer* "to propose"); second, as noted above, these speech acts are sometimes mitigated, contrary to much scholarship on commissives; third, offers do not necessarily have to be unsolicited to count as such; and finally, offers can be formed using a variety of syntactic means, such as direct questions, conditional constructions, the use of alerters and the insertion of tag questions. This final point is echoed by Searle (1976: 17) and Kissine (2009: s132), namely that no specific syntactic mood or grammatical form is inextricably linked with commissive force, as is the case with directives, for instance.

The act of refusing has formed the basis for a number of studies on commissives in Spanish. Vanderveken (1990: 185) notes that "a refusal is the illocutionary denegation of the acceptance of a request", which leads to this act's classification as a commissive. García (1992, 1999, 2007) has put forth a series of studies on refusals

in three varieties of modern Spanish: Peruvian, Venezuelan and Argentinian. In all three investigations, role plays served as the means of data collection; participants were to decline an invitation extended by an interlocutor. Results from the three studies evidence remarkably similar patterns; in all three papers, the author notes a cross-cultural paradigm for this speech act, which involves two stages. After the initial invitation, the participant responds by declining the invitation (as per the role play instructions); following the refusal, a second stage commences in which the inviter issues an insistence on the hearer's acceptance of the invitation, which is followed by a response to the insistence. All three studies code the results in terms of politeness theory, specifically whether participants invoke negative politeness (termed 'deference politeness' by the author) strategies or positive politeness ('solidarity politeness') strategies in their refusals. Both deference and solidarity strategies are evident in the data from all three groups, both in the head act of the commissive as well as the supportive moves which accompanied it. The specific distribution of strategy preferences differed somewhat between the three groups; while all three groups preferred the use of deference politeness in the first stage of the interaction, Peruvians and Argentinians (García 1992 and 2007, respectively) employed more solidarity politeness in the second stage, while Venezuelans (García 1999) showed an overwhelming preference for deference strategies in the second stage as well. In some cases, gender differences also shone through; Venezuelan men (García 1999) used deference strategies much more frequently, while Venezuelan women emphasized friendship over respect through their higher use of solidarity. García (1992) made note of an additional important cultural curiosity: in the second stage of the interactions, 65% of her Peruvian participants ended up accepting the invitation, albeit in a conditional manner, in spite of the fact that they had been instructed to decline it. The author suggests that "both the insistence by the friend and these acceptances, however nominal, are reflections of politeness in this particular socio-cultural setting" (García 1992: 237). Specifically, Peruvians appear to feel that the insistence of an invitation in the second stage as a potentially face threatening act is preferable to having the invitation sound insincere, just as nominally accepting an invitation after an insistence to do so, and thus threatening one's own negative face, is preferable to threatening the inviter's positive face with a snub.

An additional pair of studies on refusals have been put forth by Félix-Brasdefer (2003, 2006). The first of these (Félix-Brasdefer 2003) deals with the speech act of declining an invitation. The issue at hand in this investigation is that of socio-pragmatic competence as it applies to L2 learners' production of this speech act. Data for the study were collected via open role plays from both control groups of native speakers of both Spanish and English as well as the experimental group (L1 English, L2 Spanish). The results of the study show evidence of language transfer

on the part of the experimental group, both in terms of the variety of refusal strategies they choose as well as the frequency with which they employ distinct strategies. A significant finding is that American English speakers tend to be more direct than Spanish speakers when refusing invitations, particularly in situations of unequal power – those which Scollon & Scollon (2001) would label as ‘hierarchical politeness systems’. On the other hand, Latin American Spanish speakers more frequently employ mitigation strategies and avoid direct refusals with all interlocutors other than intimates, such as friends. The L2 Spanish speakers in the study evinced an intermediate level of directness in their refusals; in spite of these learners’ advanced level of language competence, the author comments that their unfamiliarity with native-like speech act behavior “suggests a lack of L2 sociocultural knowledge, indicating that even advanced learners of Spanish, who had previously lived in an L2 setting, were not aware of the L2 sociocultural values” (Félix-Brasdefer 2003: 247).

The second study (Félix-Brasdefer 2006) focuses on refusals among Mexican males in their native language. Once again, role plays were employed for data collection; these presented a number of distinct scenarios to participants in which the variables of power and social distance were manipulated. In solidarity systems (i.e., with those of equal status, such as friends), participants preferred to use direct refusal strategies, while responses became more indirect as the level of social distance increased. The author also notes that, for this Mexican community, insistence of an invitation is seen as a manifestation of the speaker’s involvement with his interlocutor and is an expected cultural behavior (2006: 2177). This finding falls in line with all three of García’s investigations on refusals noted above. Taken together, these results suggest that this type of insistence may constitute a pragmatic universal in the Spanish-speaking world.

Finally, the speech act of promising serves as the focus of a number of studies on commissives in Spanish. Boix Jovaní (2012), which takes as its corpus the medieval epic poem *El cantar de mio Cid*, notes that promises are used throughout the epic, in many cases to create expectations in the audience. Since readers know in advance that the hero will triumph in the end, the enjoyment in the tale comes in the form of seeing how the hero accomplishes this. In essence, promises in the *Cid* are reflective of the structure and psychology of medieval Spanish society – namely, that the protagonist is a model citizen of this time period and reflects such through his fulfillment of his commitments.

Rall (1993) discusses promises and their relation to cultural conceptualizations of time in Mexican Spanish. The accounts of language usage in this study are largely based on participant observation and include numerous personal accounts of cultural idiosyncrasies noted by the author. The issue of temporal deixis is at the heart of this study; Rall contends that the notion of time itself is a relative concept

and a thorough pragmatic competence in a given culture's use of commissives necessarily involves an understanding of its conception of time.

Rall's investigation raises a number of points central to the study of this category of speech acts. First, the author supports the view that commissives (specifically, promises) occur in every language and culture; in other words, they are pragmatic universals (1993: 3). Kissine (2009: 132) echoes this position by stating that "the universality of deontic commitment in human societies ... makes the dependence of commissive speech acts on a specific, intra-cultural collective convention very unlikely". However, this unlikely situation is precisely what some authors delineate for certain non-Western languages. Agyekum (2004) describes certain oaths (*ntam*) in the Akan language; the author contends that commissives as described by Searle focus strongly on individual sincerity, while Akan *ntam* commissives "include a complicit audience that shares responsibility for the fulfillment of the act to which the utterer of *ntam* commits", due to the shared moral commitments of the community (321–322). Rosaldo (1982) adds that the idea of individual commitment to carrying out a promised action is somewhat foreign to the Ilongot people and lacks corresponding speech act verbs; she elaborates that the closest equivalents to Searle's description of commissives in Ilongot either refer to the mere repetition of an action or invoke external or supernatural law, rather than individual human efforts. Thus, it is clear that the very concept of what is involved in the issuing of a commissive is highly culture-dependent.

The second issue raised by Rall is that commissives are subject to a double interpretation of sorts. She argues that, since time is a relative concept and the future is always uncertain, in the moment of the utterance of the speech act, promises are an expression of complete cooperation with the hearer, but it must be understood that commissives leave open a range of possible outcomes in the future, not only the so-called successful fulfillment of the promise (1993: 13). This philosophical point has also been widely debated in the literature. Searle (1976: 4) explains that the expressed psychological state in the performance of an illocutionary act, known as the sincerity condition, is for commissives the intention to carry out the expressed action. Tiersma (1986: 197) points out that psychological insincerity does not invalidate a commissive, since sincerity is a regulative rule of this category of speech acts and does not define them. He continues that it is not the intent to promise that commits a speaker to a certain course of action, but rather the locutionary act itself. In spite of this philosophical 'out', many authors contend that the non-fulfillment of the action expressed by a commissive can have dire consequences in real life language usage. Hardin (2001: 60) indicates that the relative infrequency of commissives in Spanish language advertising may be due to the fact that the use of these speech acts forces advertisers to be accountable for the veracity of the content of their ads. Agyekum



(2004: 321) remarks that, in Akan, commissives serve as objective guarantees of the promised action, as non-fulfillment of the task would leave the speaker open to societal prosecution. In medieval Spain, Boix Jovaní (2012: 4) explains, men who did not fulfill promises faced extreme social persecution, as they were seen as cowards who lacked honor.

Finally, and most relevant to the current study, Rall notes that promises have never been clear cut in the history of language; these speech acts have gone through marked diachronic changes in many languages (1993: 4). Culpeper & Archer (2008: 58) echo this sentiment; they argue that syntactic and semantic changes that occur in languages over lengthy periods of time can lead modern scholars to incorrectly interpret speech acts in historical varieties of language. Since we cannot assume that commissives behave in the same manner in different languages or cultures, or even at different moments in the evolution of the same language, the pragmatic employment of these acts must be traced to ensure that our modern understanding of the speech act behavior of another time is not blurred or influenced by our own. The current paper tackles one aspect of this task – namely, to synchronically document and examine commissive speech acts in 18th century Spanish as spoken in the Louisiana territory in light of the understanding of the pragmatic norms of modern Spanish outlined in this section.

### 3. Description of corpus and method

The corpus chosen for the current study consists of a collection of 100 Spanish language business letters composed in the Louisiana territory during the latter half of the 18th century.<sup>1</sup> These letters were written by various high-level government officials, including several governors of the territory during this period. The letters focus around the establishment and operations of an early Spanish settlement known as Galveztown (*Villa de Gálvez*), which was located near present-day Ascension Parish. The settlement, which was inaugurated in the late 1770s by order of the Spanish Crown, involved a significant portion of the nearly 2,400 Canary Islanders (*isleños*) who emigrated from the Spanish islands during this decade to populate and cultivate various Louisiana settlements (Din 1988: 17). Thus, this study adds to the literature on contact varieties of Spanish involving migrations (see both Philip and Ramírez-Cruz, this volume).

---

1. This collection, known as the *Galveztown Papers* collection, is housed in the archives of Louisiana State University. The original documents are held in the *Archivo General de Indias* in Seville, Spain.

The content of the letters centers on daily happenings in the settlement: problems with specific settlers; potential threats by adversaries from other nations; inventories of food and supplies, and the like. As noted by King (2011: 265), a significant portion of the content of these letters serves to petition the various governors of the territory to provide supplies and services needed by the Galveztown settlement. Unlike letters in previous centuries, particularly those sent across the Atlantic Ocean from inhabitants of the New World back to their relatives and friends in Spain, the letters examined in the present corpus did not pass through a lengthy period between their composition and their delivery. In fact, one letter in the corpus indicates that mail was carried away by messengers as soon as it was composed and was delivered in as little as two and a half hours. This extremely short transmission time facilitated the fluid communication and flow of supplies between the Galveztown settlement and the government seat in New Orleans.

Due to the nature of the letters in this collection, commissives form a second highly abundant speech act type, in addition to directives, as noted above. Given that this category of speech acts has been relatively unexplored in pre-modern language varieties and ignored in Colonial Louisiana Spanish, the current study seeks to shed light on the nature of commissives in this period. The aims of this paper, then, are threefold: first, to identify the commissive speech acts in the epistolary described here as per the classification system presented in Bilbow (2002); second, to analyze the structure of the speech acts found in the corpus with regard to the variable of the power status of each author, with reference to politeness theory (particularly Brown & Levinson 1987); and finally, to compare and contrast the findings of the current study with those presented in the literature for modern Spanish, as well as modern business language in general. The focus of this investigation is the commissive head act, which will be formally analyzed here; the supportive moves to the speech act, where present, will be discussed tangentially. With regard to social variables, the only issue at hand in the current corpus is that of the power differential between the letter writer and his intended recipient; the relationship between every pair of interlocutors involves a hierarchical politeness system, as explained by Scollon & Scollon (2001). Since all letters represented in the corpus were written by adult male government officials, other sociolinguistic variables (such as age, sex and social distance) are not considered as part of the analysis herein. The question of imposition, as laid out by Brown & Levinson (1987), will be discussed in the analysis of textual examples.

A modified version of the taxonomy presented in Bilbow (2002), a study on commissives in business language, will be employed in the analysis of the speech acts found in the present study. This author reports that the commissives found in his corpus of business communication can be divided into four categories, based on two variables: first, the level of linguistically encoded directness inherent to

the speech act – whether or not the commissive is “either marked syntactically, or indicated explicitly” (Blum-Kulka 1987: 134) and expresses commitment overtly; and second, whether or not the speech act is uttered in response to some form of initiation by an interlocutor (such as a directive or an expression of a need). Thus, commissives in the current study will be classified in the following manner (Bilbow 2002: 296):

- Direct initiated (i.e., promises)
- Direct uninitiated (i.e., offers)
- Indirect initiated (i.e., commissive hints)
- Indirect uninitiated (i.e., suggestory hints)<sup>2</sup>

The definition of directness noted above (i.e., the manner in which directness, or lack thereof, is linguistically encoded in the verb form used to express a commissive head act) will be used to code the data as belonging to one of Bilbow’s categories listed here. However, this issue will receive further comment in the discussion of textual examples below, as it is acknowledged that commissives can receive distinct interpretations based on contextual factors.

A caveat is needed on the nature of the current corpus. In spite of the long-standing debate in historical linguistics over the most appropriate type of primary sources from which to collect linguistic data (Salmon 1987; Brown & Gilman 1989; Fontanella de Weinberg 1994; Taavitsainen 1995; Culpeper & Kytö 2000; Fitzmaurice 2002; Bentivoglio 2003; King 2010), it must be recognized that written documents can be studied and appreciated *sui generis*, without reference to their potential approximation to spoken language. The discussion in the following section, and particularly the comparisons drawn between the results of the current study and those found in studies of modern Spanish, highlights the distinctions between spoken and written language and their implications for the study of speech acts.

## 4. Results

Table 1 shows the distribution of commissive head act strategies found in the corpus. Note that each token represents a unique commissive act (i.e., an act of commitment represented by a distinct verb indicating the action to be undertaken).

---

2. Bilbow (2002: 298) notes that indirect uninitiated commissives, in spite of their classification as suggestions, carry commissive and not directive illocutionary force, as in *Why don’t we look into this issue further?* This commissive subclass is discussed further in §4.

**Table 1.** Commissive head act strategies (raw frequencies)

	Inferiors	Superiors	Totals
Promises	4	8	12 (14.5%)
Offers	67	2	69 (83.1%)
Commissive hints	1	0	1 (1.2%)
Suggestory hints	1	0	1 (1.2%)
<b>Total</b>	<b>73</b>	<b>10</b>	<b>83 (100%)</b>

It can be seen from Table 1 that offers are by far the most frequent commissive strategy type in the corpus. Although they are not the exclusive strategy used by inferiors when making commitments, they comprise over 91% of the commissives used by inferiors when addressing their superiors. On the other hand, although superiors issue relatively few commissives in the corpus, their preference is clear: they issue promises in all but two isolated instances. In other words, superiors in this corpus (usually the various Spanish governors of Louisiana in the late 18th century) issue promises in response to some prodding by the officials who run the Galveztown settlement and make requests for supplies or assistance.

#### 4.1 Examples of commissive strategies

In example (1), an instance of a promise is seen:<sup>3</sup>

- (1) *Todo quanto me encarga V.S. en las Ynstrucciones **procurare ejecutarlo** al pie de la letra, y en caso que tenga alguna duda **se la participare** á V.S. afin que me aclare de ella.*

“With everything Your Lordship charges me in your instructions **I shall be certain to carry it out** to the letter, and in the event I am unsure of anything, **I will inform** Your Lordship so that you may clarify it for me.”

(22 January 1779, Francisco Collell to Bernardo de Gálvez)

Both of the commissive acts in (1) share a common grammatical form, specifically the synthetic (morphological) future tense. A distinction between the analytic future structure (*Voy a hablar* “I’m going to talk”) and the synthetic future structure (*Hablaré* “I will talk”) has been drawn in many studies on the Spanish language (cf. Kyzar, this volume). Sedano (2007: 132) notes that the synthetic future structure finds its roots in the Latin periphrasis ‘infinitival verb + auxiliary

3. All translations of corpus examples are the author’s own.

*habēre*’, which in that language implied obligation. King (2010: 262) points out that the use of the synthetic future tense to communicate directive speech acts was highly common in the Early Modern period in Spanish; a number of authors have made note of this use of the synthetic future in modern varieties of the language as well (Fleischman 1982; Gennari 2000; Hurley 1995; Márquez Reiter 2002). This combination of future time reference and a modal interpretation denoting obligation was strongly at work in Colonial Louisiana; this form is used nearly to the exclusion of the analytic future tense in the present corpus, although it is used most frequently to express obligation on the part of the speaker, not the hearer. Of the instances of commissives noted in Table 1, 89% involve the use of the synthetic future tense in the formation of the head act. Of these, nearly all instances of self-obligation were produced by inferiors; superiors appeared to prefer other linguistic forms to express commissives, as will be seen below. It could be hypothesized, then, that in this period, the synthetic future tense was recognized as the most appropriate form with which inferiors could issue commissives to their superiors.

Example (1), taken from a letter written by a commandant in charge of the Galveztown settlement, includes two instances of promises. In this section of Collell’s letter, he refers back to previous correspondence that had been sent to him by Bernardo de Gálvez, in which the governor passed along a set of instructions for overseeing the construction of a number of homes within the settlement. In the section cited in (1), Collell pledges to follow the governor’s instructions very precisely. The language used here, as is the case in many of Collell’s letters, is highly deferential; not only does he use the respectful address term *Vuestra Señoría* ‘Your Lordship’, which was appropriate in communication addressed to a high-level political leader at the time, but he also makes it clear from the content of the two commissives noted here that he intends to take whatever action is necessary to please his superior.

Example (2) presents another promise taken from the corpus:

- (2) *Apruebo el gasto de los diez pesos de una mesa y tabernaculo para la capilla, é igual cantidad para el Tablado del quarto del cirujano, por lo que hase del trabajo que vm me dice estan haciendo los negros ... me remito a lo que le escriba a vm sobre el particular Mr. Maxent, con quien he quedado acorde sobre el asunto.*

‘I approve the expense of the 10 pesos for a table and tabernacle for the chapel, and the same amount for the floorboards for the surgeon’s room. With regard to what you are doing with the work that you tell me the black settlers are carrying out ... I abide by what Mr. Maxent has written to Your Grace about the matter, as I have come to an agreement with him on this issue.’  
(12 May 1779, Bernardo de Gálvez to Francisco Collell)

The commissive patterns exhibited in this letter, written by Governor Gálvez, closely pattern after those in example (1). In this portion of his letter, Gálvez responds to a series of requests on the part of Collell, who has asked for the governor to approve funding for a variety of materials and services needed within the settlement. The primary distinction between the commissives in the two letters is a grammatical one: the governor formulates both of his commissives here (*apruevo* and *me remito*) using the present tense. Unlike with inferiors, there is no general linguistic pattern followed by superiors in the corpus in the expression of commissive head acts: while some make use of the synthetic future, as in (1), a greater number employ the present tense, as in (2), and others utilize indirect formulations to express self-commitment. It is acknowledged, however, that the lack of a consistent pattern of commissive behavior on the part of superiors cannot be as clearly established as that seen for inferiors, due to the paucity of data from these speakers found in the corpus.

With regard to the content of the commissives expressed in this example, it is clear that (2) is another instantiation of a promise (as defined in §3), as its content is expressed directly, yet is initiated by requests made by Collell.<sup>4</sup> This type of commissive head act is employed primarily by superiors in the corpus, as seen in Table 1. This leads to the potential conclusion that superiors in this time period typically did not make uninitiated commitments, but rather responded with this speech act type when prodded to do so. This result is in concordance with what we would expect to happen in this circumstance, according to Brown & Levinson's (1987) theory of face wants. In the interest of protecting one's own negative face,

---

4. Example (2) presents an interesting counterexample to the theoretical notions regarding commissives often raised in the literature. As noted by a reviewer, *aprovar* (modern *aprobar*) "to approve" does not obligatorily refer to a future action, and it often involves an action that will be undertaken by a third party and not the speaker. In addition, speech acts of this type are frequently regarded as declaratives and not commissives (cf. Vanderveken 1990). However, I argue that, in the context of the letters in the current corpus, the speech act introduced by *apruevo* in (2) indeed satisfies the requirements for commissives listed herein for two reasons. First, as mentioned in §4, this speech act is produced by Gálvez in response to a request on the part of Collell for the governor to provide services and goods. Thus, the speech act is initiated by a prior directive, which is the second criterion of a promise as noted in §3. Second, with the uttering of the verb "I approve", the governor is effectively accepting the request and promising, albeit indirectly, to provide funding for the goods for which Collell has petitioned. In other words, what is being promised is the provision of funds, something the governor himself enacts, and not an action on the part of a third party. Although it is not the case here, it has been noted in the literature that commissives can, and frequently are, made on behalf of, or involving action of, someone other than the speaker (cf. Bilbow 2002; Danet 1980; Hardin 2001, others).

speakers will attempt to avoid being tied down by commitments to the greatest extent possible so as to maintain freedom of action. While this is often difficult for inferiors to accomplish, by virtue of their position, superiors can avoid making self-commitments if they wish to do so. Bilbow (2002: 298) remarks that some speech acts which are intended as commissives by speakers of lower status can be heard as directives when uttered by those of high status, due to nothing more than the level of power and position of the speaker. Thus, the speech act of promising reveals significant distinctions between the linguistic behaviors of superiors and inferiors in the present corpus.

In the following example, an instance of an offer is given:

- (3) *me ara Vm<sup>d</sup> el favor de entregarle Mil y quinientos pesos en papeles de Rey los mas chicos que Vm<sup>d</sup> pueda pues quiero retirar todos los papeles mios que tengo en este puesto, que a fin de Agosto le entregare a Vm<sup>d</sup> todas las quantas y quedaremos claro del todo le ymbio yncluzo el Recivo de dha. cantidad para su resguardo.*

“Your Grace will do me the favor of paying him 1,500 pesos in the King’s currency, the smallest bills that you have, as I want to take away all of the documents that I have in this place. At the end of August I will submit all of the bills to Your Grace and we will then be clear on everything. I am also sending you a receipt for said amount for your records.”

(7 April 1784, Antonio Maxent to Juan Ventura Morales)

The text in (3) presents an interesting counterexample to the patterns determined by power status seen in the majority of the corpus. Antonio Maxent, a wealthy French planter, was appointed Lieutenant Governor of the Louisiana province in 1779 by Governor Bernardo de Gálvez, who happened to be his son-in-law. Two years later, in October 1781, King Charles III of Spain elevated Maxent to the position of “Lieutenant-Governor and Captain-General in all matters relating to the respective Indian nations that inhabit the provinces of Louisiana and West Florida” (Watson 1996: 311). Thus, at the time of writing of the letter in (3), Maxent held what was arguably the position of second greatest power in Louisiana. Juan Ventura Morales, on the other hand, served as Governor Gálvez’s secretary during his military campaigns in 1779 and 1781, and later he held a variety of lower-level positions (such as accountant and legal advisor) in the colony. Morales’ most notable service came in 1796, when he took over the position of intendant of Louisiana. According to all accounts,<sup>5</sup> Morales, an incredibly headstrong man, had constant conflicts with both Governors Casa-Calvo and Gayoso, particularly the latter. These conflicts centered around the fact that, a decade prior to Morales’

---

5. See, for instance, Holmes (1996) and Montero de Pedro (2000).



taking over the office (in December 1786), King Charles decreed that the position of intendant would no longer be subordinate to any governor – effectively limiting the powers of the governor of Louisiana and expanding those of the intendant of the province. Thus, by the last four years of the eighteenth century, Morales held what could be seen to be the most powerful (or certainly the second most powerful) position in the Spanish government of Louisiana.

At the time of composition of the letter in (3), however, Morales was over a decade away from taking his post as intendant. In 1784, with Maxent holding the position of Lieutenant Governor and Captain General, his position clearly gave him power over Morales, who was still acting in a variety of capacities such as secretary and accountant. The letters in the corpus between these two men focus almost exclusively on matters of record keeping and the repayment of debts within the colony. The content of (3) fits in with this pattern; this letter focuses entirely on matters of bookkeeping and the assignment and repayment of debts. What makes the letters between Maxent and Morales captivating is the seeming linguistic reversal of power roles. It can be seen in (3) that Maxent addresses Morales as *Vuestra Merced*, the address form appropriate between most men in what Scollon & Scollon (2001) would term relationships involving deference politeness, as well as for superiors when writing to inferiors in this period (cf. example (2)). However, his language when writing Morales goes far beyond “politic behavior” (Watts 2003): in many places he uses deferential language and even speaks of “gladly serving him [Morales] in any way I can” (*en todo quanto le pueda servir que lo hare con mucho gusto*, 8 November 1781).

In (3), we see an example of an offer; Maxent pledges to send Morales information on debt accounts involving Galveztown settlers, in spite of the fact that this action was not requested by Morales. Given the information presented in relation to (2) above, namely that superiors typically did not make uninitiated commitments to inferiors in this time period, we may raise the question as to why Maxent would do so. The answer may be found in an issue of interpersonal relations. It is generally acknowledged in the historical literature<sup>6</sup> that Morales was intensely unpopular in Louisiana, due to his “dictatorial and arrogant temperament” and his extreme zeal in “the faithful and punctual fulfillment of his duties” (Montero de Pedro 2000: 114); however, in Madrid, he was supported as a highly faithful and effective public servant. Thus, it is conceivable that Maxent, in wishing to avoid Morales’ well-known wrath, chose to adopt a deferential stance in his communications with Morales, in spite of his higher power status. In Brown & Levinson’s (1987) terms, not only does Maxent exhibit considerable negative

---

6. See Din (1996), Holmes (1996), and Montero de Pedro (2000).



politeness towards Morales via his repeated employment of deferential language, but he goes above and beyond what might be expected and exhibits positive politeness as well through his (possibly exaggerated) conveying of cooperation with the intendant-to-be.

An offer of a distinct nature is seen in (4):

- (4) *Como es por el favorable amparo de V.S. que conseguí primariamente la comandancia de esta villa ... espero con el mismo merezer la gracia de que V.S. tenga á bien de honrrarme y escogirme en el numero de los demas, para acompañar a V.S. en la expedicion de Pensacola, baxo la Ynteligencia que por el cervicio de S.M. y el particular de V.S. deramare hasta la ultima gota de mi sangre.*

“Given that I first obtained command of this village through the positive recommendation of Your Lordship ... I hope to again merit this same grace in that Your Lordship will see fit to honor me and choose me among those who will accompany Your Lordship in the expedition to Pensacola. Under His Majesty’s intelligence, and in his service, as well as that of Your Lordship, I will spill the last drop of my blood.”

(14 June 1780, Francisco Collell to Bernardo de Gálvez)

In this letter, Collell writes Governor Gálvez to request that he be permitted to leave his post at Galveztown in order to accompany the governor and a group of officials on an expedition to Florida. As Collell is prone to do throughout the corpus, he fills his letter with highly dramatic statements of his commitment to Gálvez and the settlement. By the standards of the time and the nature of the correspondence involved, the language Collell uses is highly deferential, to the point of becoming flowery; he includes numerous exaggerated examples of positive politeness to make his cooperation with the governor’s wishes clear (*Siendo asi que la voluntad de V.S. es la mia, de qualquiera manera que V.S. disponga de mi persona quedare satisfecho* “Being that Your Lordship’s will is also mine, in any way that Your Lordship sees fit to make use of me, I will be satisfied”). In (4), another such instance, taken from the beginning of this letter, is seen: Collell expresses his loyalty to the governor as well as to the king of Spain via an unsolicited blood oath (*deramare hasta la ultima gota de mi sangre*). Although one has trouble imagining that Governor Gálvez would ever call upon Collell to carry out the terms of this oath, the example remains as a commissive speech act, specifically an offer, under the terms of the present analysis.

Example (5) displays a commissive hint:

- (5) *suplico á V.S. se digne de embiarme algunos Ynstrumentos de Guerra, afin que por ellos, y (ayudado de la poca esperiencia que tengo) haga respetar la vandera de mi Rey, y la persona de V.S., y Prometo de defender ambos sentidos hasta perder la ultima gota de mi sangre ... en caso que sea nesesario de hazer un fuerte mediano no me faltara Yngenio para ello ...*

“I beg that Your Lordship deign to send me some instruments of war, so that with them (aided by the little experience I have) I can make others respect the flag of my king, as well as Your Lordship, and **I promise to defend** both senses until I lose my last drop of blood ... in the event that it becomes necessary to construct a medium-sized fort, **I will not be lacking in ingenuity** to do so ...”

(16 February 1779, Francisco Collell to Bernardo de Gálvez)

In this letter, we see one of the few corpus examples of a commissive speech act which involves a performative verb (*apruebo* “I approve” in example (2) above constitutes another of these). The first commissive in (5) employs the verb *prometer* “to promise”, which, according to the taxonomy used in the current study, is interestingly categorized as an offer, given that it is an uninitiated, yet directly stated, commissive. Once again, we see Collell’s use of grandiloquent language at work; for the second time in the corpus, we note the commandant’s voluntarily offered blood oath.

The second commissive in this letter, however, is even more notable: when discussing the possibility of the need to construct a fort to protect the Galveztown settlement from potential attacks of the British, Collell asserts to Governor Gálvez *no me faltara Yngenio para ello* “I will not be lacking in ingenuity to do so”. This commissive is stated indirectly, without any overt reference to the action in question noted in the head act, but it has evidently been initiated on the part of the governor in a prior letter or conversation. This is the sole corpus example of what Bilbow (2002) terms a commissive hint. The current finding is somewhat at odds with Bilbow’s result that 13% of the commissives in his corpus were coded as this subtype. The author notes that, in the data from his study on intercultural business meetings, commissive hints “are usually highly modalized and accompanied by prosodic features that indicate a speculative reaction and a low level of commitment” (2002: 297). Clearly, this definition cannot be applied to the data in the present study, given that they are taken from written documents and thus cannot be analyzed in terms of their prosody. It is possible, however, that the second half of the definition applies well to the current example; Collell engages in speculation when discussing the proposed fort, given that, from the context of this letter, it is apparent that no formal plans for the fort have been made at the time of writing. It is difficult to judge Collell’s level of commitment to this project from a written document; however, it can be noted that this is the only reference to this endeavor in the current letter, and the only commissive he issues related to it is expressed indirectly – as opposed to the commandant’s overt use of a performative verb earlier in the same letter, as well as his euphuistic language style.

Finally, (6) presents an instance of a suggestory hint:

- (6) *Abisare a V.S. si es cierto que los Yngleses tengan un camino que pase por frente de este fuerte ... para venir a esta banda tienen que pasar dos leguas de lago, en donde si V.S. halla por conveniente se puede poner un destacamento de 4 hombres para vigilar, y dar aviso á esta en caso de alguna novedad.*

“I will notify Your Lordship if it is true that the English have a path which passes in front of this fort ... in order to get to this side they have to pass through two leagues of lake, where, if Your Lordship feels it is appropriate, a detachment of four men **could be placed** in order to keep watch and give warning to this side in case anything comes up.”

(26 September 1779, Francisco Collell to Bernardo de Gálvez)

This section of Collell’s letter begins with an offer, namely to provide information to Governor Gálvez once he has investigated the landscape in question in preparation for a possible attack by the British. Following this, the commandant makes a suggestion to the governor regarding a way in which the (now built) fort at Galveztown could be further protected. Bilbow (2002: 298) states that suggestory hints, such as the second commissive in (6), are “the most ambiguous type of commissive speech act”, due to four characteristics that they share: first, they are uninitiated by another speech act; second, they often refer to actions that are not well-defined; third, they do not specify whether the agent of the action is to be the speaker or another party; and fourth, they are frequently highly hedged. This description fits the commissive in (6) well; although Collell does note the involvement of a detachment of four men, he does not specify exactly what the men would do. Unlike the majority of the commissives he issues (as seen in several examples above), in this case he does not offer to personally take on the action specified in the head act – although it could be assumed that the commandant would likely be the person in charge of such an action, as the official who runs the Galveztown settlement. The commissive head act itself, *se puede poner un destacamento de 4 hombres para vigilar* “a detachment of four men could be placed in order to keep watch” makes use of the Spanish ‘impersonal *se*’ construction; the syntax used here is ambiguous and could be understood as purely impersonal (“one could place a detachment of four men”) or as an example of the passive voice, as seen in the translation above. In either case, Collell deliberately avoids mentioning the agent of the proposed action and further hedges the head act with the phrase *si V.S. halla por conveniente* “if Your Lordship feels it is appropriate”. This type of hedging is seen in Brown & Levinson’s model (1987: 145) as a negative politeness strategy, by which the speaker attempts to avoid coercing the hearer. Similar to Brown & Levinson, Blum-Kulka et al. (1989: 288) analyze the type of hedge in (6) as a type of supportive move to a speech act which they term an “imposition

minimizer". This indirect, uninitiated and somewhat vague commissive stands in stark contrast to Collell's open, self-sacrificing style seen in nearly every letter he pens throughout the corpus.

## 5. Discussion and conclusions

As was noted in §3, the goals of the current study are threefold: first, to identify the commissive speech acts in the corpus of letters under study as per the taxonomy presented in Bilbow (2002); second, to analyze the structure of the commissives found in the corpus with regard to the variable of power; and third, to compare the findings of the current study with those presented in similar studies on modern Spanish, as well as modern business language in general.

The distribution of commissive speech acts in this corpus, given in Table 1, bears witness to the overwhelming preference for direct yet uninitiated commissives (i.e., offers). This commissive strategy type is used nearly to the exclusion of all others by inferiors when writing to their superiors in the epistolary. This finding suggests that, in business communications in colonial Louisiana, it was considered appropriate – perhaps even compulsory – for subordinates to take the initiative in offering their services in various ways to their superiors. It is important to note, however, that, in spite of lesser amounts of data in the corpus representing their language use, superiors display the opposite tendency – namely, to issue commissives only after being prompted to do so by a subordinate requesting goods or assistance. Bilbow's (2002) results present the reverse distribution: 46% of the data collected in his study were classified as promises, as compared to offers, which comprised 24% of the commissives tabulated. However, as this author did not specifically classify instantiations of commissives according to the power status of the speaker, this aspect of his investigation cannot be directly compared to the current study.

It has also been seen here that both subtypes of commissives classified as indirect formulations (commissive and suggestory hints) are nearly absent from the epistolary. Again, this finding contrasts with Bilbow (2002); in this author's study, nearly 30% of the commissives analyzed were classified as indirect. This finding raises an important question with regard to the nature of the corpora involved in these two investigations: while Bilbow's study focuses on the use of spoken language in business meetings, the current study employs as its corpus a collection of business letters. It is foreseeable that the authors of business letters would revert to indirect strategies less frequently than those engaged in real-time, face-to-face communication; unlike with face-to-face communication, recipients of letters are unable to rely on paralinguistic cues such as gestures and body language to assist

them in interpreting the force and reference of a speech act. As was noted in §4.1, Bilbow makes reference to features such as the use of nods of the head and prosodic cues to indicate speculation being common in his corpus; these features are unavailable to letter writers and recipients in the process of transmitting messages. Not only do these types of paralinguistic information clarify the force of a speech act, but they also serve to mitigate potential face threats present in the issuing of distinct speech acts. Authors of letters must rely solely on syntactic and lexical means to accomplish this goal. It is also evident that, as the possibility of immediate feedback and clarification is not possible when reading a letter (at least, this was not a possibility in 18th century Louisiana), letter writers are aware of the need to express their thoughts as clearly and directly as possible, particularly when urgent matters are at hand. This factor alone explains the highly infrequent occurrences of indirect commissive formulations in the present corpus. This finding has been echoed in a number of other studies on both familiar and business letters in pre-modern language varieties (for instance, King 2010, 2011).

With regard to the relation of the present study to investigations on modern varieties of Spanish, a number of points can be noted. It must first be acknowledged that, due to the relative paucity of studies on commissives in modern Spanish, coupled with the small number of commissive tokens found for superiors in the present corpus (see §2 and §4 above), it would be irresponsible to claim that any definitive diachronic conclusions could be drawn from the present study. What can be described, however, are general trends found in the 18th century data and their relation to what has been reported for modern varieties of the language. The current study concurs with many of the conclusions reached by Chodorowska-Pilch (2002) for offers in modern Peninsular Spanish – namely, that commissives are rarely expressed via performative verbs; that these speech acts are sometimes mitigated; and that commissives are expressed via a variety of syntactic means. These issues were all raised and discussed in §4.1 in relation to the linguistic examples provided. Second, the current study differs from the finding of Hardin (2001) in that commissives in her corpus were frequently qualified or hedged. As discussed in relation to example (6) above, hedged commissives are rare in the current corpus and stand in stark contrast to the generally direct style employed in the majority of the epistolary. As mentioned in §2, the hedged commissives mentioned by Hardin can be considered marked and dispreferred forms, as hedging only serves to weaken the obligation inherent in the speech act. This could serve as an effective strategy for advertisers, as in Hardin's corpus, but would beg the question of sincerity if seen in significant numbers in the current study. Finally, with regard to the issue of diachronic changes in speech act behavior raised by Rall (1993), the current study has revealed a crucial issue – namely, the recurrent use of the synthetic future tense in the formation of commissives. It

appears clear that in 18th century Louisiana, the morphological future form was employed nearly ubiquitously in the expression of commissive force, particularly by subordinates when communicating with their superiors. As was seen in §4.1, the use of the synthetic future has been noted in a number of studies on directive speech acts in both modern and pre-modern varieties of Spanish, but has rarely, if ever, been discussed to date in such studies on commissives.

The present study not only adds to the literature on language use in Colonial Louisiana Spanish, but it also augments the limited literature on commissive speech acts in this language. Future investigations can add to the body of knowledge on this topic by looking at other sparsely studied eras and varieties of Spanish, as well as other types of primary sources (familiar letters, literary works, etc.) to show the range of linguistic behavior inherent to speech act production. Although the current corpus is large enough to show cogent trends from the community in question, a broader set of data sources would allow for a more thorough look into commissive behavior in this period. It is hoped that future studies take up this task and expand our knowledge of diachronic speech act behavior.

## Acknowledgements

I would like to thank the staff of the LSU Libraries' Special Collections, and in particular Judy Bolton, for their assistance with this project.

## References

- Agyekum, Kofi. 2004. "Ntam 'Reminiscential Oath' Taboo in Akan". *Language in Society* 33:3.317–342. doi:10.1017/S0047404504043015
- Astington, Janet W. 1988. "Children's Production of Commissive Speech Acts". *Journal of Child Language* 15:2.411–423. doi:10.1017/S0305000900012423
- Austin, John L. 1962. *How to Do Things with Words*. Oxford: Oxford University Press.
- Bentivoglio, Paola. 2003. "Spanish Forms of Address in the Sixteenth Century". *Diachronic Perspectives on Address Term Systems* ed. by Irma Taavitsainen & Andreas Jucker, 177–191. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/pbns.107.09ben
- Bernaerts, Lars. 2010. "Interactions in *Cuckoo's Nest*: Elements of a Narrative Speech-Act Analysis". *Narrative* 18:3.276–299. doi:10.1353/nar.2010.0004
- Bernicot, Josie & Virginie Laval. 1996. "Promises in French Children: Comprehension and Metapragmatic Knowledge". *Journal of Pragmatics* 25.101–122. doi:10.1016/0378-2166(94)00075-5
- Bilbow, Grahame T. 2002. "Commissive Speech Act Use in Intercultural Business Meetings". *International Review of Applied Linguistics* 40.287–303. doi:10.1515/iral.2002.014

- Blum-Kulka, Shoshana. 1987. "Indirectness and Politeness in Requests: Same or Different?" *Journal of Pragmatics* 11.131–146. doi:10.1016/0378-2166(87)90192-5
- Blum-Kulka, Shoshana, Juliane House & Gabriele Kasper, eds. 1989. *Cross-Cultural Pragmatics: Requests and Apologies*. Norwood, New Jersey: Ablex.
- Boix Jovaní, Alfonso. 2012. "Promesas y juramentos en el *Cantar de Mio Cid*". *Bulletin of Hispanic Studies* 89:1.1–13. doi:10.3828/bhs.2012.1
- Brewer, Talbot M. 2003. "Two Kinds of Commitments (and Two Kinds of Social Groups)". *Philosophy and Phenomenological Research* 66:3.554–583. doi:10.1111/j.1933-1592.2003.tb00279.x
- Brown, Roger & Albert Gilman. 1989. "Politeness Theory and Shakespeare's Four Major Tragedies". *Language in Society* 18.159–212. doi:10.1017/S0047404500013464
- Brown, Penelope & Stephen C. Levinson. 1987. *Politeness: Some Universals in Language Usage*. Cambridge: Cambridge University Press.
- Bruder, Kurt A. 1998. "A Pragmatics for Human Relationship with the Divine: An Examination of the Monastic Blessing Sequence". *Journal of Pragmatics* 29.463–491. doi:10.1016/S0378-2166(97)00062-3
- Byers, A. Martin. 1999. "Intentionality, Symbolic Pragmatics, and Material Culture: Revisiting Binford's View of the Old Copper Complex". *American Antiquity* 64:2.265–287. doi:10.2307/2694278
- Carvalho, Vitor R. & William W. Cohen. 2005. "On the Collective Classification of Email 'Speech Acts'". *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval* ed. by Gary Marchionini et al., 345–352. New York: Association for Computing Machinery.
- Chodorowska-Pilch, Marianna. 2002. "Las ofertas y la cortesía en español peninsular". *Actos de habla y cortesía* ed. by María E. Placencia & Diana Bravo, 21–36. Munich: Lincom.
- Csordas, Thomas J. 1997. "Prophecy and the Performance of Metaphor". *American Anthropologist* 99:2.321–332. doi:10.1525/aa.1997.99.2.321
- Culpeper, Jonathan & Dawn Archer. 2008. "Requests and Directness in Early Modern English Trial Proceedings and Play Texts, 1640–1760". *Speech Acts in the History of English* ed. by Andreas Jucker & Irma Taavitsainen, 45–84. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/pbns.176.05cul
- Culpeper, Jonathan & Merja Kytö. 2000. "Data in Historical Pragmatics: Spoken Interaction (Re)cast as Writing". *Journal of Historical Pragmatics* 1:2.175–199. doi:10.1075/jhp.1.2.03cul
- Danet, Brenda. 1980. "Language in the Legal Process". *Law & Society Review* 14:3.445–564. doi:10.2307/3053192
- DeJong, Huib M. & Wouter G. Werner. 1998. "Continuity and Change in Legal Positivism". *Law and Philosophy* 17:3.233–250.
- Desilva, David A. 1994. "Despising Shame: A Cultural-Anthropological Investigation of the Epistle to the Hebrews". *Journal of Biblical Literature* 113:3.439–461. doi:10.2307/3266781
- Din, Gilbert C. 1988. *The Canary Islanders of Louisiana*. Baton Rouge: Louisiana State University Press.
- Din, Gilbert C. 1996. "Spain's Immigration Policy in Louisiana and the American Penetration, 1792–1803". *The Louisiana Purchase Bicentennial Series in Louisiana History*, vol. 2, *The Spanish Presence in Louisiana 1763–1803* ed. by Gilbert C. Din, 334–351. Lafayette, Louisiana: Center for Louisiana Studies of the University of Southwestern Louisiana.
- Dippold, Doris. 2007. "Using Speech Frames to Research Interlanguage Pragmatics: Facework Strategies in L2 German Argument". *Journal of Applied Linguistics* 4:3.285–308.



- Félix-Brasdefer, J. César. 2003. "Declining an Invitation: A Cross-Cultural Study of Pragmatic Strategies in American English and Latin American Spanish". *Multilingua* 22.225–255. doi:10.1515/mult.2003.012
- Félix-Brasdefer, J. César. 2006. "Linguistic Politeness in Mexico: Refusal Strategies among Male Speakers of Mexican Spanish". *Journal of Pragmatics* 38.2158–2187. doi:10.1016/j.pragma.2006.05.004
- Fitzmaurice, Susan M. 2002. *The Familiar Letter in Early Modern English*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/pbns.95
- Fleischman, Suzanne. 1982. *The Future in Thought and Language: Diachronic Evidence from Romance*. Cambridge: Cambridge University Press.
- Fletcher, Judith. 2003. "Women and Oaths in Euripides". *Theatre Journal* 55:1.29–44. doi:10.1353/tj.2003.0020
- Fontanella de Weinberg, María B. 1994. "Fórmulas de tratamiento en el español americano (siglos XVI y XVII)". *El español en el nuevo mundo* ed. by María B. Fontanella de Weinberg, 7–31. Washington: Organización de los Estados Americanos.
- Fraser, Bruce & William Nolen. 1981. "The Association of Deference with Linguistic Form". *International Journal of the Sociology of Language* 27.93–109.
- Galveztown Papers. n.d. Ms. 890, Louisiana and Lower Mississippi Valley Collections. Baton Rouge: LSU Libraries.
- García, Carmen. 1992. "Refusing an Invitation: A Case of Peruvian Style". *Hispanic Linguistics* 5:1–2.207–243.
- García, Carmen. 1999. "The Three Stages of Venezuelan Invitations and Responses". *Multilingua* 18:4.391–433. doi:10.1515/mult.1999.18.4.391
- García, Carmen. 2007. "'Ché, mirá, vos sabés que no voy a poder': How Argentineans Refuse an Invitation". *Hispania* 90:3.551–564. doi:10.2307/20063564
- Gennari, Silvia. 2000. "Semantics and Pragmatics of Future Tenses in Spanish". *Hispanic Linguistics at the Turn of the Millennium* ed. by Héctor Campos et al., 264–281. Somerville, Mass.: Cascadilla.
- Hancher, Michael. 1979. "The Classification of Cooperative Illocutionary Acts". *Language in Society* 8:1.1–14. doi:10.1017/S0047404500005911
- Hardin, Karol J. 2001. *Pragmatics of Persuasive Discourse in Spanish Television Advertising*. Arlington, Texas: SIL International / University of Texas at Arlington.
- Haugh, Michael. 2008. "Intention in Pragmatics". *Intercultural Pragmatics* 5:2.99–110. doi:10.1515/IP.2008.006
- Haverkate, Henk. 1994. *La cortesía verbal: Estudio pragmatolingüístico*. Madrid: Gredos.
- Holmes, Jack D. L. 1996. "The Provincial Governor-General: Manuel Gayoso de Lemos". *The Louisiana Purchase Bicentennial Series in Louisiana History*, vol. 2, *The Spanish Presence in Louisiana 1763–1803* ed. by Gilbert C. Din, 94–112. Lafayette, Louisiana: Center for Louisiana Studies of the University of Southwestern Louisiana.
- Holtgraves, Thomas. 2005. "The Production and Perception of Implicit Performatives". *Journal of Pragmatics* 37.2024–2043. doi:10.1016/j.pragma.2005.03.005
- Huang, Yan. 2010. "Speech Acts". *Concise Encyclopedia of Philosophy of Language and Linguistics* ed. by Alex Barber & Robert J. Stainton, 705–714. Oxford: Elsevier.
- Hurley, Joni Kay. 1995. "Pragmatics in a Language Contact Situation: Verb Forms Used in Requests in Ecuadorian Spanish." *Hispanic Linguistics* 6–7.225–264.
- Johansson, Ingvar. 2003. "Performatives and Antiperformatives". *Linguistics and Philosophy* 26:6.661–702. doi:10.1023/B:LING.0000004557.12602.2C



- King, Jeremy. 2010. "Pleas from the New World: The Structure of Directive Head Acts in Colonial Spanish Familiar Letters". *Journal of Historical Pragmatics* 11:2.250–276. doi:10.1075/jhp.11.2.04kin
- King, Jeremy. 2011. "Power and Indirectness in Business Correspondence: Petitions in Colonial Louisiana Spanish". *Journal of Politeness Research* 7.259–283. doi:10.1515/jplr.2011.013
- Kissine, Mikhail. 2009. "Illocutionary Forces and What Is Said". *Mind & Language* 24:1.122–138. doi:10.1111/j.1468-0017.2008.01356.x
- Knops, Andrew. 2006. "Delivering Deliberation's Emancipatory Potential". *Political Theory* 34:5.594–623. doi:10.1177/0090591706290780
- Larson, Gerald James. 1978. "Prolegomenon to a Theory of Religion". *Journal of the American Academy of Religion* 46:4.443–463. doi:10.1093/jaarel/XLVI.4.443
- Leech, Geoffrey. 1983. *Principles of Pragmatics*. London: Longman.
- Liu, Si. 2011. "An Experimental Study of the Classification and Recognition of Chinese Speech Acts". *Journal of Pragmatics* 43.1801–1817. doi:10.1016/j.pragma.2010.10.031
- Márquez Reiter, Rosina. 2002. "A Contrastive Study of Conventional Indirectness in Spanish: Evidence from Peninsular and Uruguayan Spanish". *Pragmatics* 12:2.135–151. doi:10.1075/prag.12.2.02mar
- Márquez Reiter, Rosina & María Elena Placencia. 2005. *Spanish Pragmatics*. Hampshire: Palgrave Macmillan. doi:10.1057/9780230505018
- McCallam, David. 2003. "The Nature of Libertine Promises in Laclos's *Les liaisons dangereuses*". *The Modern Language Review* 98:4.857–869. doi:10.2307/3737929
- McDonald, Lee C. 1969. "Myth, Politics and Political Science". *The Western Political Quarterly* 22:1.141–150. doi:10.2307/446153
- Montero de Pedro, José. 2000. *The Spanish in New Orleans and Louisiana*. Trans. Richard E. Chandler. Gretna, Louisiana: Pelican.
- Myhill, John. 1996. "The Development of the Strong Obligation System in American English." *American Speech* 71:4.339–388. doi:10.2307/455712
- Natsopoulos, Demetrios, Grouios, George, Bostantzopoulou, Sevasti, Mentenopoulos, Georges, Katsarou, Zoe & Jim Logothetis 1993. "Algorithmic and Heuristic Strategies in Comprehension of Complement Clauses by Patients with Parkinson's Disease". *Neuropsychologia* 31:9.951–964. doi:10.1016/0028-3932(93)90150-X
- Olson, David R. 2007. "Self-Ascription of Intention: Responsibility, Obligation and Self-Control". *Synthese* 159:2.297–314. doi:10.1007/s11229-007-9209-2
- Pérez Hernández, Lorena. 2001. "The Directive-Commissive Continuum". *Miscelánea* 23.77–98.
- Pérez Hernández, Lorena & Francisco José Ruiz de Mendoza. 2002. "Grounding, Semantic Motivation, and Conceptual Interaction in Indirect Directive Speech Acts". *Journal of Pragmatics* 34.259–284. doi:10.1016/S0378-2166(02)80002-9
- Person, Raymond F. 2009. "'Oh' in Shakespeare: A Conversation Analytic Approach". *Journal of Historical Pragmatics* 10:1.84–107. doi:10.1075/jhp.10.1.05per
- Polansky, Susan G. 1988. "Textual Coherence in the Duke of Rivas's *El desengaño en un sueño*: The Dramaturgy of Destiny". *Modern Language Studies* 18:3.3–17. doi:10.2307/3194964
- Rall, Marlene. 1993. "'Mañana te hablo': La deixis temporal en el acto de la promesa y su entorno cultural". *Discurso* 15.1–15.
- Rosaldo, Michelle Z. 1982. "The Things We Do with Words: Ilongot Speech Acts and Speech Act Theory in Linguistics". *Language in Society* 11:2.203–237. doi:10.1017/S004740450009209
- Rudanko, Juhani. 2004. "'I wol sterve': Negotiating the Issue of a Lady's Consent in Chaucer's Poetry". *Journal of Historical Pragmatics* 5:1.137–158. doi:10.1075/jhp.5.1.07rud

- Salmon, Vivian. 1987. "Sentence Structures in Colloquial Shakespearian English". *A Reader in the Language of Shakespearean Drama* ed. by Vivian Salmon & Edwina Burness, 265–300. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/sihols.35.29sal
- Sánchez, Marta E. 1980. "Inter-Sexual and Intertextual Codes in the Poetry of Bernice Zamora". *Multi-Ethnic Literature of the United States* 7:3.55–68. doi:10.2307/467028
- Sbisà, Marina. 2001. "Illocutionary Force and Degrees of Strength in Language Use". *Journal of Pragmatics* 33.1791–1814. doi:10.1016/S0378-2166(00)00060-6
- Scollon, Ron & Suzanne Wong Scollon. 2001. *Intercultural Communication: A Discourse Approach*. Malden, Mass.: Blackwell.
- Searle, John R. 1976. "A Classification of Illocutionary Acts". *Language in Society* 5:1.1–23. doi:10.1017/S0047404500006837
- Sedano, Mercedes. 2007. "Future Tense Expressions in Several Spanish Corpora". *Working with Spanish Corpora* ed. by Giovanni Parodi, 132–144. London: Continuum.
- Sitter, Stefan & Adelheit Stein. 1992. "Modeling the Illocutionary Aspects of Information-Seeking Dialogues". *Information Processing & Management* 28:2.165–180. doi:10.1016/0306-4573(92)90044-Z
- Taavitsainen, Irma. 1995. "Interjections in Early Modern English: From Imitation of Spoken to Conventions of Written Language". *Historical Pragmatics: Pragmatic Developments in the History of English* ed. by Andreas H. Jucker, 439–465. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/pbns.35.23taa
- Tiersma, Peter M. 1986. "The Language of Offer and Acceptance: Speech Acts and the Question of Intent". *California Law Review* 74:1.189–232. doi:10.2307/3480357
- Tiersma, Peter M. 1993. "Linguistic Issues in the Law". *Language* 69:1.113–137. doi:10.2307/416418
- Trosborg, Anna. 1995. "Statutes and Contracts: An Analysis of Legal Speech Acts in the English Language of the Law". *Journal of Pragmatics* 23.31–53. doi:10.1016/0378-2166(94)00034-C
- Vanderveken, Daniel. 1990. *Meaning and Speech Acts*, vol. 1. Cambridge: Cambridge University Press.
- Verschuere, Jef. 1983. "Review of *Speech Act Classification: A Study in The Lexical Analysis of English Speech Activity Verbs*". *Language* 59:1.166–175. doi:10.2307/414064
- Visconti, Jacqueline. 2009. "Speech Acts in Legal Language: Introduction". *Journal of Pragmatics* 41.393–400. doi:10.1016/j.pragma.2008.06.007
- Watson, Thomas D. 1996. "A Scheme Gone Awry: Bernardo de Gálvez, Gilberto Antonio de St. Maxent, and the Southern Indian Trade". *The Louisiana Purchase Bicentennial Series in Louisiana History*, vol. 2, *The Spanish Presence in Louisiana 1763–1803* ed. by Gilbert C. Din, 307–316. Lafayette, Louisiana: Center for Louisiana Studies of the University of Southwestern Louisiana.
- Watts, Richard J. 2003. *Politeness*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511615184
- Wisse, Maarten. 2000. "The Meaning of the Authority of the Bible". *Religious Studies* 36:4.473–487. doi:10.1017/S0034412500005400



## CHAPTER 7

# Discourse markers in variation

## *Por tanto* vs. *por lo tanto*

Sarah Sinnott

University of Sioux Falls

This chapter explores the variation between the Spanish discourse markers *por tanto* and *por lo tanto* from a pragmatic perspective. The data, extracted from a corpus built using blog commentary, are categorized according to the type of causality expressed in the construction, as well as whether or not the speaker is involved in a verbal argument. Analysis reveals that *por lo tanto* is used more frequently than *por tanto* when epistemic causality is expressed as well as when the speaker is engaged in an argument. This leads to the conclusion that *por lo tanto* contributes greater argumentative strength to an utterance than does *por tanto*.

### 1. Introduction

In this paper I set out to account for the pragmatic variation between the Spanish discourse markers (DMs) *por lo tanto* (PLT) and *por tanto* (PT), shown in examples (1)–(2), specifically in Peninsular Spanish. These DMs are not distinguished from one another in grammars of the Spanish language. In some cases they are lumped together, as seen in Martín Zorraquino & Portolés Lázaro (1999: 4001), who discuss PT and state that the form “cuenta con la variante *por lo tanto*” (exists along with the variant *por lo tanto*). In other reference grammars only one of the two is mentioned, for example Butt & Benjamin (2000: 459), according to whom PLT is used to “show that what follows is the result of what preceded”. They do not specifically mention PT. Considering the recent growth in interest in studies of DMs (many of which will be discussed in this article) in conjunction with the study of language variation in general, one must assume that the existing habit of grouping together these two markers is not an accurate description of their function. In what follows, I explore the uses of PT and PLT in context in order to

determine what differentiates them from each other and therefore why one might be used over the other.

- (1) *Pedí la epidural y me la pusieron mal **por tanto** sentía la mitad del cuerpo y en consecuencia dolor.* (MR)  
 “I asked for an epidural and they did it wrong **por tanto** I could feel half of my body and so I felt pain.”
- (2) *Un negocio es tiempo y dinero. **Por lo tanto** es normal que se opongan a lo que para ellos no es rentable.*  
 “A business is time and money. **Por lo tanto** it’s normal that they oppose anything that is not profitable to them.” (MR)

I show that the two variants differ in terms of the type of causality expressed in utterances in which they are used. Specifically, PLT occurs more frequently than PT when relationships of epistemic causality are expressed. In addition, PT and PLT differ in terms of their argumentative strength. PLT contributes higher argumentative strength than PT.

In order to demonstrate this claim, I begin with an overview of existing literature on the DMs themselves, followed by descriptions of causality and argumentation theory. I then discuss the methodology chosen for obtaining and analyzing tokens of the connectors. Finally, I explore the relationship between each variant, causality type and argumentation.

## 2. Previous literature

There is a general lack of research regarding these DMs specifically and, as mentioned above, they are typically grouped together under the same definition or description. Discourse markers, in general, are used to indicate the type of relationship that exists between elements of a discourse; their content is not truth conditional (Blakemore 2004). According to many investigators, DMs have two roles. Travis (2005) claims that one of these roles is to define the interpretation of an utterance in relation to other utterances; the second is to express any attitude that the speaker has toward the content of the utterance. Portolés Lázaro (1993) agrees with this dual description, claiming that DMs have a grammatical function, which serves to provide instructions for the hearer as to the relationship between discourse segments, and a pragmatic function, which may give rise to a number of inferential meanings. Fraser (2009) refers to this “grammatical function” as procedural meaning, which is one of his three characteristics of a DM. The other characteristics are that the marker must be lexical and that it must be hosted by the second segment of the discourse elements it connects.

We will see the basic nature of the procedural meaning of PT and PLT in this section. The remainder of the chapter guides us toward a more complete understanding of this and their additional pragmatic functions (in Portolés Lázaro's terms). In other words, this chapter addresses the type of relationship between discourse elements that PT and PLT express, as well as some additional inferences connected to their use.

Fraser (2009: 9) would most likely classify PT and PLT as 'inferential DMs', which "signal ... that S1 provides a basis for inferring S2" (S = segment) (see Brody, this volume, for further discussion of Fraser's classification of DMs). Portolés Lázaro (1993) classifies them as 'consecutive connectives', indicating that the content of the latter element (consequent) is reinforced by the former (antecedent).

Because there are not many available descriptions of these connectives, I give the definition from the *Diccionario de la lengua española* (2014) for each connective in (3)–(4). The definition of PLT is more detailed than that of PT. The fuller definition of PLT indicates that the consequent was motivated by what was said in the antecedent. In addition, the definitions provide other phrases or DMs that indicate a similar relationship between antecedent and consequent.

- (3) *por tanto: loc. adv. Por lo que, en atención a lo cual.*  
"por tanto: locution. adverb. Because of that, regarding that."
- (4) *por lo tanto: loc. adv. Por consiguiente, por lo que antes se ha dicho, por el motivo o las razones de que acaba de hablarse.*  
"por lo tanto: locution. adverb. As a result, because of what was said prior, because of the causes or reasons that were just said."

As noted, Martín Zorraquino & Portolés Lázaro (1999: 4100) offer one description of PT and PLT combined in the *Gramática de la lengua española*. They state that the markers introduce "un consecuente que se obtiene después de un razonamiento a partir de ... antecedente" (a consequent that is obtained through a process of reasoning based on the antecedent). In addition, they claim that these markers may be infelicitous when content causation (to be described below) between events is expressed because there is no reasoning process involved.

In his 2001 book, Portolés Lázaro only refers to PT, and classifies it as an adverb (called 'commentary pragmatic markers' by Fraser 2009). He points out, however, the features that distinguish it from other adverbs. For example, it is non-variable in form, it cannot occur alone (5) and its content cannot be questioned, negated or paraphrased (6) (examples from Portolés Lázaro 2001).

- (5) *¿Por qué lo harás?*  
"Why will you do it?"  
\*Por tanto.

- (6) *Y por tanto no iré a la fiesta.*  
 “And *por tanto* I will not go to the party.”  
 #¿*Por tanto?* ¿*De verdad?*  
 “#*Por tanto?* Really?”

According to Montolío (2001: 101), PT and PLT “indica(n) la conclusión que se deduce de la información previa” (indicates that the conclusion is deduced from the previous information). This study does not distinguish between the two variants either, stating that “*por (lo) tanto* ... puede aparecer con o sin el elemento anafórico *lo-* de hecho, lo más común es que se omita” (*por (lo) tanto* ... can appear with or without the anaphoric *lo-* in fact, it is more common that it is omitted).

Martínez (1997: 49) claims that the connectives “Se antepone(n) a una oración que expresa una consecuencia de lo que antes se ha dicho o la conclusión a que se llega” (are utilized before a phrase that expresses a consequence that is deduced from the previous information). While she claims that the DM can be used “indistintamente con anafórico o no” (with or without the anaphoric *lo*), she adds that PLT “dota de cierto <<rigor>> a la consecuencia que se deriva” (adds a little ‘something’ to the derived consequent) (p. 48, emphasis in the original). While this statement does not offer an in-depth explanation, it is, to my knowledge, the only reference to a possible difference in usage between the two variants. Being that this reference is contained within a writing style guide and not a linguistic analysis, we do not expect to see any deeper analysis, however this intuition is not far adrift from my hypothesis that the markers differ in terms of their argumentative strength.

Other than Montolío’s claim already noted that PT is more commonly used than PLT, no information is provided in the existing documentation of the distribution of the two markers. A search of several existing corpora reveals the distributions seen in Table 1. The first three corpora have similar distributions; however, the trend is reversed in the fourth. This difference may be due to the fact that the COREC radio portion is exclusively oral, whereas the data found in the other corpora is not necessarily so. In addition, some of these corpora are diachronic in nature. At this point we do not know what changes, if any, may have occurred over time in the use of these two markers, and therefore any comparisons made must be done with caution.

All of the descriptions presented here are in agreement that the procedural meaning of PT and PLT tells the hearer that the consequent was reasoned from the antecedent. We do not have any accounts of any possible pragmatic functions to work with. The question, then, is what motivates the use of one form over the other. I provide evidence to support Martínez’s intuition that PLT does differ from PT. I show that the two differ in terms of the type of causality they express and that their relative argumentative strength differs. PLT is used more frequently to express epistemic causality while imposing greater argumentative strength than PT.

**Table 1.** Frequencies of *por tanto* and *por lo tanto* in various corpora (expressed as percentages of occurrence)

	PT	PLT
Corpus del español (Davies 2002) 100,000,000 words	.7	.3
CORDE (Real Academia Española 2014) 250,000,000 words	.64	.36
CREA (Real Academia Española 2014) 170,000,000 words	.74	.26
CORLEC (radio portion) (Marcos Marín 1992) 510,200 words	.4	.6

## 2.1 Causality

According to Sweetser (1990), there are three types of causality that are conceived of by the mind and that can be expressed with language. These are content, epistemic and speech act causality. At this time, I only consider content and epistemic causality. When *p* and *q* are related via content causality, *p* or *q* actually causes the other in a real-world way. Epistemic causality, on the other hand, describes a relationship in which the knowledge expressed in *p* or *q* is used as a basis for the conclusion expressed in the other clause.

In some languages, such as English, one DM can be used to express multiple types of causality. Such is the case with *because*, which can be used to express all three of Sweetser's causality types. In other cases, a DM might be used to express fewer types of causal relations. Such is the case with the connective *since*, which is only used to express epistemic or speech act causality but not content causality (Sweetser 1990).

In the case of PT and PLT it is the antecedent that causes the consequent, either via content or epistemic causality. In example (1) above, the misadministration of the epidural caused the patient to be able to feel her body. This is an example of content causality because the occurrence of one real-world event (administering the epidural incorrectly) caused another event to occur (the speaker's ability to feel). In example (2), the knowledge that businesses require time and money caused the speaker to conclude that it is normal for business owners to oppose anything that is not profitable to their business. This is an example of epistemic causality because the speaker arrived at her conclusion after a reasoning process stemming from existing knowledge. In other words, "the consequence is not some state of affairs in the spatio-temporal world, but a mental state of the protagonist, i.e., the concluder" (Pander Maat & Degand 2002: 220).

According to Noordman & de Blijzer (2000), epistemic relations are more indirect than content relations. This claim is based on the idea that even epistemic conclusions rely on the knowledge of real-world cause and effect. Considering example (2) then, Noordman & de Blijzer might say that the epistemic



conclusion that it is normal for business owners to object to the policies can only be reached if one understands that the real-world event of needing time and money could cause the real-world event of opposing certain policies that might restrict that. Their hypothesis that epistemic relations were more indirect was supported when they found that these relations required more time to process than did content relations.

Another difference between epistemic and content relations is described by Pander Maat & Degand (2002). They state that epistemic conclusions can be thought of as being “weak” or “faulty” (p. 222). Content consequents, on the other hand, can only be judged by virtue of having happened or not; either the consequent occurred or it did not. For example, looking back at (1), there is potential that the woman might be lying about being able to feel her body, but not whether it was the ‘correct’ result of the poor administration of the epidural. We simply take it for granted that it was the result. We do not respond, for example, by saying something like “Well, you didn’t necessarily feel anything”. In (2), however, we could claim that the conclusion that it is normal for business owners to object to the policies is a faulty conclusion and we might have a different conclusion in mind. Here we might respond with something like “Well, it’s not normal because ...”.

Fraser (2009) suggests that further research into the functioning of DMs within Sweetser’s content, epistemic and speech act domains is needed and this chapter answers that call with empirical data demonstrating the function of PT and PLT within two of these domains. I argue that PLT is more likely than PT to connect a consequent that is arrived at via epistemic reasoning. In order to determine whether PT and PLT do in fact differ in terms of the type of causality they coincide with, we will need to evaluate cases in which the markers occur and determine what type of causation is involved in each. Before discussing how this was done, let us explore argumentation theory, which has been used to account for the behavior of other DMs and will play a role in this analysis.

## 2.2 Argumentation theory

Anscombe & Ducrot (1994) claim that all utterances have an argumentative orientation that guides the discourse toward coming utterances. Any given utterance therefore opens the way for utterances to follow. This, in turn, makes some continuations felicitous while others are not. These authors insist that any semantic description of a language feature include any relevant information about its argumentative orientation. This information has been included to some degree in recent descriptions of DMs, including several in Spanish. According to Portolés Lázaro (2001), there are DMs, such as *además* “in addition”, which are used to connect co-oriented utterances. Co-oriented utterances are utterances that

lead towards the same conclusion. An utterance that might be co-oriented with Portolés' example of *Manuel ha bebido un poco* might be *no tiene sus gafas*. Both utterances lead to the conclusion that Manuel should not drive. Because they are co-oriented, they could be connected by *además*, such as seen in (7).

- (7) *Manuel ha bebido un poco y, además, no tiene sus gafas.*  
 “Manuel has had a little to drink and, *además*, he doesn't have his glasses.”

Other DMs are used to connect anti-oriented utterances, or utterances that do not guide toward the same conclusion. Reig Alamillo (2011) describes the DM *lo que pasa es que* “what happens is that” and claims that this marker is used to introduce an argument that orients toward a different conclusion than the previous. In (8), from Reig Alamillo (2011), the speaker claims that he feels like going out to eat. This has an argumentative orientation toward a conclusion that he go out to eat. However, by employing *lo que pasa es que*, the speaker is able to offer an anti-oriented utterance that supports the conclusion of not going out to eat.

- (8) *Me apetece salir a cenar, lo que pasa es que tengo que trabajar.*  
 “I feel like going out for dinner, *lo que pasa es que* I have to work”

There are also DMs (including *pero* “but”) that can introduce conclusions that are unexpected in the sense that the preceding utterance does not orient to them. One such DM is *sin embargo* “nevertheless”. As seen in (9), from Portolés Lázaro (2001), the first clause does not naturally orient to the conclusion in the second clause, but the sequence is logical due to the use of *sin embargo*.

- (9) *Alicia es inteligente. Sin embargo, le suspenden los exámenes.*  
 “Alicia is smart. *Sin embargo*, she failed her exams.”

Another important component of Anscombe & Ducrot's theory of argumentation is the proposal that two utterances that orient to the same conclusion exist on a scale of the relative strength with which they lead to that conclusion. In addition, when a higher element on the scale is used, any member below it could be felicitously used as well; however, the reverse is not true. Portolés Lázaro (2005) provides the Spanish example in (10).

- |  |  |
|--|--|
| (10) Antecedent<br>+ <i>Fuerza</i><br>“strength”<br>– <i>Es joven.</i><br>“He's young.”<br>– <i>No es viejo</i><br>“He isn't old.” | Consequent<br><br><i>Tiene toda una vida por delante.</i><br>“He has his whole life ahead of him.” |
|--|--|

In this case, both *es joven* and *no es viejo* can lead to the conclusion that *tiene toda una vida por delante*. However, *es joven* leads to this conclusion with greater argumentative strength. In essence, the jump from *es joven* to *tiene toda una vida por delante* is easier than the jump from *no es viejo* to *tiene toda una vida por delante*. The consequent is more negotiable or more subjective when the argument is *no es viejo*, than when it is *es joven*. Finally, in any context in which *es joven* is used, *no es viejo* could also be used. However, the use of *no es viejo* does not entail the felicitous use of *es joven*.

The DM *es más* “what’s more” is used to connect a segment to a co-oriented segment of greater argumentative strength (Portolés Lázaro 2011). Utilizing the same arguments from the scale above, we might say something like (11), but not (12), because *es más* cannot introduce a weaker member of the scale.

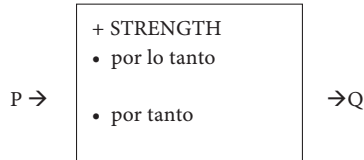
- (11) *No es viejo, es más, es joven.*  
 “He’s not old, **es más**, he’s young.”
- (12) *#Es joven, es más, no es viejo.*  
 “#He’s young, **es más**, he’s not old.”

Portolés Lázaro (2001) also describes the DM *con todo* “even so” in terms of argumentative strength. In this case, the marker is used to connect an antecedent with an anti-oriented consequent. In (13), from Portolés Lázaro (2001), Luisa being hardworking does not ordinarily lead to the conclusion that she will not get her work done on time. The use of *con todo* allows for this unexpected consequent. *Con todo* differs from connectors such as *pero* “but”, however, in that it adds the information that the antecedent is very highly anti-oriented to the given consequent, which, in turn, makes the consequent more surprising. *Con todo* therefore adds argumentative strength to the antecedent, making it more anti-oriented to the consequent than it might seem without the use of the marker.

- (13) *Luisa es extremadamente trabajadora. Con todo, no conseguirá presentar el informe a tiempo.*  
 “Luisa is extremely hardworking. **Con todo**, she won’t manage to present the report on time.”

As we have seen, PLT and PT guide the discourse by indicating that the information in q can be deduced from the information in p. Therefore, it can be said that that these DMs link an argument in p which orients toward the conclusion in q. I will argue in this chapter that the two markers differ in that PLT adds argumentative strength that would not be there if PT were used to connect the clauses. Therefore, p orients more strongly toward q when PLT is used to connect them than when PT is used to connect them. So, PLT and PT can be placed on a scale similar to that seen in (10). This scale is shown in Figure 1. This scale demonstrates that

the argumentative strength with which *p* orients to *q* is greater when PLT is used than when PT is used. As such, a speaker might choose to use PLT when he or she wishes to strengthen the relationship between the two segments, thereby making the conclusion seem more definitive. This would be in line with Martinez's claim (see above) that PLT “dota de cierto <<rigor >>” (1997: 48).



**Figure 1.** Argumentative strength of *por lo tanto* vs. *por tanto*

If PT and PLT do differ in terms of their argumentative strength, we might expect PLT to be used in cases in which a speaker has extra motivation to convince an interlocutor that the conclusion he or she is proposing is the correct conclusion to be drawn from the argument or antecedent presented. To test this hypothesis, I evaluate which variant is used in cases in which the speaker is actually arguing against another person. In these cases a speaker, rather than simply stating his or her point of view, is generally trying to convince another that his point of view is the correct one. I explain how this evaluation was carried out in the next section.

### 3. Methodology

The data used in this study are from a corpus compiled from five different blogs featured on the Spanish website *20minutos.es*. This is a website produced in Spain, and the website has another version for its Latin American audience; therefore, all bloggers and commenters were assumed to be speakers of Peninsular Spanish (unless evidence suggested otherwise, as described below). There has been recent growth in the use of social media for compiling corpora for a variety of reasons, including register and quantity, for example (see Riccelli, this volume for further discussion and another example of a corpus built using social media). The five blogs utilized all deal with different subject matter and included *Madre reciente* (a blog about motherhood and children), *El ojo de Gran hermano* (about the popular reality show *Gran hermano*), *La gente de Rosy Runrún* (a gossip blog about famous people), *Ni libre ni ocupado* (written by a taxi driver) and *Bravo Fernando* (a blog about Formula 1 racing). In this chapter I discuss data extracted from the first three of these blogs. Unless otherwise indicated, all examples and data presented in this paper are from one of these blogs. The blog is indicated with the following abbreviations following

each example: GH = *Gran hermano*, MR = *Madre reciente*, and RR = *La gente de Rosy Runrún*. In addition, all data points are presented exactly as written in the original blog post and, therefore, due to the informal nature of the medium, the orthography may not conform to standardized conventions.

All tokens of PT and PLT were extracted from these blogs and their commentary using a Google site search. The results of each search were saved using EverNote software. Tokens were discarded if they were spam commentary, quoted from another source or written by an author who explicitly stated that he or she was not Spanish.

No personal information on any of the participants in these blogs was obtained. The only information associated directly with each data point is the screen name of its author. These screen names are not presented with the data. No effort was made to ascertain information on sex, age or any other characteristic based on the author's screen name or the content or grammar of their post. Hence, all data points are anonymous.

For the purposes of this study, I only take into account tokens found in the commentary following each blog entry. Tokens in the blog entries themselves have not been considered at this time in order to avoid possible undue speaker effects on the variation of PT and PLT. Within the blog commentary itself, there were very few instances of multiple comments posted by the same user name. Each token was coded for a variety of factors. Those of most interest to this study include the type of causality involved, with the categories being content or epistemic, and whether or not PT/PLT occurred as part of an argument against the blog post or another commenter. Six ambiguous cases (in which it could not be determined whether the causality type was epistemic or content) were excluded from the results. There were 493 tokens included in this study, making it larger than other quantitative studies of DMs. For example, Pander Maat & Sanders (2000) used a corpus containing 50 tokens each of three Dutch causal connectives. Travis (2005) based her conclusions on a more robust sample of 605 tokens divided among four DMs.

In order to determine the causal relation between *p* and *q* in each token, each example was analyzed as to whether real-world events were being described or whether conclusions were being made by the commenter. Travis (2005) states that it is very difficult to determine the exact scope of DMs in many cases. Lagerwerf (1998) tells us that often times an utterance taken out of context can be ambiguous with regard to the causal relation expressed. Because of these issues, the entire comment was used as context, as well as surrounding comments or the text of the blog itself if needed.

An example of content causality from the corpus is seen in (1) above. An example of epistemic causality from the corpus is seen in (2) above. I provide two

more examples of each causality type in (14)–(16) below. In (14), we see a case of content causality in which not being used to drinking soymilk causes intolerance to it. In (15) we have an example of epistemic causality in which the speaker's idea that all of the players (on *Gran hermano*) are all the same causes him to conclude that it will not be any more unfair for Ivan to win than it would be for anyone else to win. In (16), both causality types are expressed in one comment. The first is epistemic; the speaker concludes that emotions are biological because he doesn't believe in anything spiritual. The second is content; the good feelings that oxytocin creates cause dependencies.

- (14) *Demonizamos la leche y endiosamos la soja, que es un producto al que no estamos acostumbrados y **por tanto** nuestro nivel de tolerancia es mucho menor.* (MR)  
 “We demonize milk and glorify soy, which is a product that we're not used to and **por tanto** we have less tolerance for it.”
- (15) *Y si Telecinco ha decidido que el que gane sea Ivan pues que gane Ivan que vale lo mismo que los demás y **por lo tanto** no es tanta injusticia.* (GH)  
 “And if Telecinco has decided that Ivan is the one who is going to win, then let him win. He's the same as the others and **por lo tanto** it's not really unfair.”
- (16) *Yo no creo en lo espiritual, **por lo tanto** las emociones tiene un origen neurológico o bioquímico ... La oxitocina genera bienestar, y **por tanto** una “dependencia” de la persona que relacionamos con ese bienestar.* (sic) (MR)  
 “I don't believe in anything spiritual, **por lo tanto** emotions have a neurological or biochemical origin ... Oxytocin creates good feelings, and **por tanto** a ‘dependency’ on the person with whom we associate those feelings.”

Argumentation was determined in the following manner. Each comment was analyzed for signs that the commenter was arguing the point of another commenter or blog. These signs included explicit claims of contradiction, such as seen in (17), in which the speaker claims that she is the one who has the real information, not the blogger. The explicit contradiction is shown in bold typeface. I also included insults, as in (18), in which the commenter tells another blogger that she is pathetic for telling lies. The insult is given in bold.

- (17) ***yo si tengo informacion real y eres una falsa.** De que Mario esta enamorado y Maria siempre lo ha hecho madurar eso ha sido desde ya hace tiempo y **por lo tanto** Mario nunca han estado separado* (sic) (RR)  
 “**I have the real information and you are a liar.** The stuff about Mario being in love and Maria making him more mature is old news and **por lo tanto** Mario has never been separated.”

- (18) *No presumas de lo que desconoces solo por hacerte la interesante, que todos sabemos ya, porque no aciertas nada, y por lo tanto no hay nadie que te cuente lo que dices que te cuentan, que eres una mísera muerta de hambre.* (RR)

“Don’t say what you don’t know just to make yourself sound interesting, we already know that you don’t get anything right, and **por lo tanto** there isn’t anybody who tells you what you claim they tell you, **you are miserably pathetic.**”

If any of these signs were present, the argumentation was confirmed by locating the disputed content in the surrounding commentary or in the blog text. If no obvious signs of dispute were present, the token was not further explored and the token was classified as non-argumentative.

The distribution of the DMs in relation to causality type and argumentation is discussed in the following sessions.

4. Results

In the three blogs described above, there were 493 tokens: 211 of PT and 282 of PLT. This distribution in and of itself varies from some of the distributions in sources discussed above. At 43% PT and 57% PLT, it is closest to the distribution found in the radio portion of the COREC. In addition, the distribution is counter to the tendency reported by Montolío (2001) that PT is more frequent. This difference may have to do with text type, but I make no claim to that effect at this point.

I now present the results specifically with regard to causality and argumentation.

4.1 Causality

As seen in Table 2, there were a total of 216 tokens of content causality and 277 of epistemic causality. Recall Martín Zorraquino & Portolés Lázaro’s (1999) claim that neither PT nor PLT is used with content causality. The results of this study, in which 44% of the tokens represent content causality, provide counterevidence to this notion.

Table 2. Distribution of discourse markers and causality type

	Content causality	Epistemic causality	Total N
PT	114	97	211
PLT	102	180	282
Totals	216	277	493

Upon considering the distribution of the DMs within each causality type, we see that PT and PLT have approximately equal distribution within cases of content causality. PT is slightly higher with 52.8% of the content causality tokens. Within epistemic causality, however, PLT is used significantly more frequently than PT; PLT makes up 65% of the epistemic causality tokens.

Looking at the distribution of causality type within the use of each discourse variant, we see that PT appears about equally in both types of causality; 54% of PT tokens are found within content causality and 46% within epistemic. PLT shows a difference in frequency, appearing much more often in cases of epistemic causality (64%) than in content causality (36%). The distribution of DM vs. causality type is statistically significant at the .05 level ( $X^2 = 15.6358$ ;  $p = 7.7E-05$ ), meaning that causality type does indeed correlate with DM selection.

## 4.2 Argumentation

To determine the role of argumentation in these DMs, I look at the relationships between argumentation, DM, and causality type.

Table 3 shows the distribution of argumentation within these tokens. Approximately 30% of all tokens are found within arguments. 58% of the tokens found within arguments use PLT, slightly more than PT. This distribution is not statistically significant, but some interesting patterns immerge that are worthy of discussion.

**Table 3.** Distribution of discourse markers and argumentation

	Argumentation	Non-argumentation	Total N
PT	61	150	211
PLT	84	198	282
<b>Totals</b>	<b>145</b>	<b>348</b>	<b>493</b>

If we look at the distribution of causality type and argumentation (Table 4), we see that this distribution is statistically significant at the .05 level ( $X^2 = 16.7268$ ;  $p = 4.3E-05$ ). This seems reasonable, because one might assume that, since consequents of a content nature are more objective (as discussed above), they are therefore less likely to be argued. When no argument is involved, there is an approximately equal likelihood of each causality type. In other words, epistemic and content causality relations are expressed with the same frequency when the speaker is not arguing with another interlocutor. This, again, stands to reason if we consider that arguments tend to be about conclusions drawn by the speaker rather than real-world events. Once we take out the arguments, which are skewed to epistemic causality, we are left with an equal distribution of causality type.



**Table 4.** Distribution of argumentation and causality type

	Epistemic	Content	Total N
Argument	102	43	145
No argument	175	173	348
Totals	277	216	493

Additional interesting results in this area are seen when we break the numbers down further. As shown in Table 5, the distribution of PT and PLT in relation to epistemic and content causality is still significant at the .05 level when we only include the argumentation tokens ( $X^2 = 6.4778$ ;  $p = 0.010923$ ). While the distribution of DM and causality are significant both overall and among only argumentative examples, the frequencies are slightly different in the subgroup vs. the overall group.

**Table 5.** Distribution of discourse markers and causality type within argumentation

	Content causality	Epistemic causality	Total N
PT	25	36	61
PLT	18	66	84
Totals	43	102	145

In general we see a higher percentage of epistemic causality relations within the argument data when compared to the overall data at 70% vs. 56.2%. This is to be expected, because we have already discussed that arguments are more likely to be epistemic. However, this increase in epistemicity occurs across both DMs. That is, the number of epistemic relations that contain PT goes up as well as the number of epistemic relations that use PLT. Within the argumentative examples, the percentage of epistemic causality PT tokens is 24.8% while the overall percentage of epistemic causality PT tokens is 19.7%. Similarly, the percentage of epistemic causality PLT tokens is 45.5% within the argumentative examples but 36.5% overall. Again, this tendency holds regardless of the DM utilized. With this increase, of course content causal relations decrease in frequency. 17.2% of the argumentative tokens are content causality PT, while 23.1% of overall tokens are content causality PT tokens. Finally, 12.3% of argumentation tokens are content causality PLT, and 20.7% of tokens overall are of this nature.

We can also look at this from the point of view of DM frequency within causality type among the argument tokens in comparison with the whole set. Upon doing so we see a slight trend of increased use of PT within content causality at

52.8% in the overall corpus vs. 58.1% here. We also see an increase of PLT use among epistemic causality at 57% in the overall corpus vs. 65% here. In other words, PT is associated more strongly with content causality and PLT is associated more strongly with epistemic causality in the argumentation tokens than they are in the overall corpus.

In sum, the distribution of PT and PLT among arguments vs. non-arguments is not statistically significant. However, the argument data patterns similarly to the overall data in terms of the use of PT vs. PLT when expressing content or epistemic causality and the relationship is stronger than in the overall data.

## 5. Discussion

These distributions indicate that PLT occurs significantly more frequently than PT within cases of epistemic causality. This supports the hypothesis that the two connectives vary by causality type and specifically that PLT occurs more frequently to introduce consequents arrived at via reason rather than by real-world events.

In order to determine whether PLT contributes more argumentative force than PT, we looked at the use of the variants in cases of argumentation. There was no correlation between DM and argumentation, but, nonetheless, the results do support this hypothesis when taken in combination with the causality results, as I now show.

PLT was the most frequent variant in epistemic causal relations. The consequents in these relations are arrived at by reason, and therefore they may be considered faulty, as mentioned above.

PLT was again significantly more frequent than PT within argumentative epistemic causals. In these cases we have a speaker trying to convince an interlocutor not only that the other interlocutor is wrong but that the speaker's reasoned conclusion is the correct one.

PLT occurred at about the same rate as PT in the overall corpus among content relations. PLT content tokens made up 20.7% of the data, and PT content tokens made up 23.1% of the data. However, within argumentation, PLT content tokens made up only 12.4% of the data and still occurred less frequently than PT content relations, which made up 17.2% of the argumentation data.

When taken together, it appears that PLT is used most frequently when the most is at stake for the speaker, or when the speaker has the most to prove. The speaker has the burden of proof when expressing his or her reasoned conclusions. The speaker carries a burden when trying to win an argument. The speaker has an even greater burden when trying to win an argument with those reasoned conclusions.

It would appear then that PLT is utilized as a tool to lessen this burden. As such, it might signal to the interlocutor that, although the speaker's conclusion might seem like a faulty or weak one, it is not. It signals this by adding argumentative force that tells the hearer that the antecedent is stronger evidence for the consequent than might be assumed otherwise.

Two questions remain, then: why is PLT used at all in content cases, and why is PT used in epistemic cases? Regarding the first, the fact is that even events that seem to be related via real-world causality may actually involve a speaker's conclusions (Pander Maat & Degand 2002). Sanders & Spooren (2010: 23) also discuss the possibility of "subjectified information". This is information that, while fact, is presented by the speaker as if it were her own point of view. Let us look at (19) below, which was considered to be a token of content causality in this study. In such a case, yes, it might be true that the status of the women causes them to not need an amniocentesis, and this might be a rule typically followed in the medical practice. However, it also might be the case that the speaker has concluded for herself that this is the case. She has said that she knows that the women are at low risk and therefore concludes that they do not need an amniocentesis.

- (19) *La mayoría de enfermos de síndrome de dawn son por embarazos de mujeres sin riesgo (menos de 35 años) y por lo tanto que no se les obliga la amniocentesis.* (sic)

"The majority of people with Down's Syndrome are born to low-risk women (under 35 years of age) and **por lo tanto** they aren't required to have an amniocentesis."

To the second question regarding the use of PT in epistemic causal relations, not all epistemic conclusions require 'convincing'. To demonstrate, consider the example in (20), which was categorized as epistemic in this study. In fact, it was also considered to be a token of argumentation. Upon studying this utterance, we find a great clue that perhaps this speaker is not too concerned with the probability that her conclusion will be accepted by her interlocutors. She presents quite a bit of evidence that should be readily available to the public. If they accept that evidence, then the conclusion that Clara Lago is a television actress is undeniable. In addition, this speaker mitigates her argument in several ways. She includes the friendly greeting *hola a todos*, she makes clear that she only wants to clarify something and she does not accuse anyone in particular of being wrong. In sum, this speaker does not need the argumentative strength of PLT because she has undeniable evidence to support her claim and she does not appear to be trying to 'win' an argument with another interlocutor.

- (20) *Hola a todos!! Solo una puntualización, eso de que Clara Lago (estupenda actriz por cierto) no hace televisión no es del todo cierto, ha intervenido*

*en varias series tanto como personaje fijo como episódico, por ejemplo ha participado en Raquel busca su sitio, compañeros, hospital central, lex, los hombre de Paco .... También ha intervenido en el corazón del oceano (pendiente de estreno) por tanto esta chica si que hace televisión. (sic)* (RR)

“Hi everybody! Just a note, regarding Claro Lago (a great actress for sure) not being on TV, that’s not completely true. She has been in several series as a regular or as a guest actor. For example, she was in *Raquel busca su sitio*, *Compañeros*, *Hospital central*, *LEX*, *Los hombres de Paco* .... She was also in *El corazón del océano* (which hasn’t come out yet) **por tanto** this girl does do TV.”

In sum, PLT seems to be used when the most is at stake, either when the speaker is presenting an epistemic and therefore refutable conclusion as the best one or when he or she is trying to convince another that his or her conclusion is right while the other’s is wrong. The use of PLT provides extra argumentative force to make his or her conclusion more convincing.

In turn, PT seems to be used when there is nothing at stake. The speaker is either presenting facts or is not worried that (or does not care if) his or her conclusion might not be accepted as valid. In these cases, no extra argumentative strength is needed to strengthen the relationship between the antecedent and the consequent, either because it is already strong or because it does not matter.

Finally, I do not discount the potential role of subjectivity in the variation of these two forms. Upon further examination, we may find that PLT is used to add argumentative force to more subjective content relations or that PT is used to ‘objectify’ epistemic relations.

## 6. Conclusions

This study directly contributes to our understanding of the functions of the Spanish DMs *por tanto* and *por lo tanto* by demonstrating that causality type correlates with their use due to differences in the argumentative strength that each marker contributes. Specifically, PLT has the pragmatic function of adding argumentative strength to the antecedent, making it a stronger argument for the conclusion presented. Because of this, it is more frequently utilized to introduce an epistemic conclusion.

From a broader perspective, this demonstration shows that Sweetser’s theory of causality and Anscombe & Ducrot’s theory of argumentation can work together to account for language variation, and this adds to our growing knowledge of the function of DMs overall.

While this study has shed light on the role of previously unstudied DMs in Spanish, much remains to be done. As already mentioned, the role of subjectivity needs to be explored in order to determine if it is an additional factor in this variation. Additionally, here we have only taken into account the use of these forms in a pan-Peninsular manner. It remains to be seen whether these conclusions hold true in other dialects or whether there are differences within the Peninsula.

I further suggest that, in order to provide more evidence supporting the role of argumentative force and causality in the variation of these markers, perception and acquisition studies need to be undertaken. Perception studies might determine whether an interlocutor is aware of this added force or not. Acquisition studies are of interest because, as described in Sanders & Spooren (2010), Kyratzis et al. demonstrated in their 1990 study that epistemic causal relations are acquired late. If PLT is associated with epistemic causality, the assumption would be that it is acquired after PT.

In sum, these findings contribute to our understanding of the interactions between the cognitive processes that allow us to perceive and express causality and our own desire to utilize language to achieve our goals, in this case to convince others of the strength of our conclusions.

## References

- Anscombre, Jean-Claude, & Oswald Ducrot. 1994. *La argumentación en la lengua* ed. by Jean-Claude Anscombre & Oswald Ducrot; Spanish version ed. by Julia Sevilla & Marta Tordesillas. Madrid: Gredos.
- Blakemore, Diane. 2004. "Discourse Markers". *The Handbook of Pragmatics* ed. by Laurence Horn & Gregory Ward, 221–240. Mass.: Blackwell.
- Bravo, Fernando (username). n.d. "Bravo, Fernando". Retrieved from 20minutos.es.
- Butt, John & Carmen Benjamin. 2000. *A New Reference Grammar of Modern Spanish*. Chicago: NTC Publishing Group.
- Davies, Mark. 2002–. *Corpus del Español: 100 million words, 1200–1900s*. Retrieved from <http://www.corpusdelespanol.org>.
- Díaz, Daniel. n.d. "Ni libre ni ocupado". Retrieved from 20minutos.es.
- Diccionario de la lengua española. n.d. "tanto". Retrieved from <http://lema.rae.es/drae/?val=tanto>.
- Fraser, Bruce. 2009. "An Account of Discourse Markers". *International Review of Pragmatics* 1:2.293–320. doi:10.1163/187730909X12538045489818
- Hernández, Gus. n.d. "El Ojo de Gran Hermano". Retrieved from 20minutos.es.
- Lagerwerf, Luuk. 1998. *Causal Connectives Have Presuppositions: Effects on Coherence and Discourse Structure*. The Hague: Holland Academic Graphics.
- Madreciente (username). n.d. "Madre reciente," Retrieved from 20minutos.es.
- Marcos Marín, Francisco, ed. 1992. *Corpus Oral de Referencia de la Lengua Española Contemporánea: Corpus Oral Peninsular*. Retrieved from [www.lllf.uam.es/~fmarcos/informes/corpus/corpusix.html](http://www.lllf.uam.es/~fmarcos/informes/corpus/corpusix.html).

- Martín Zorraquino, María Antonia & José Portolés Lázaro. 1999. "Los marcadores del discurso". *Gramática descriptiva de la lengua española* ed. by Ignacio Bosque & Violeta Demonte, vol. 3, 4051–4213. Madrid: Espasa-Calpe.
- Martínez, Roser. 1997. *Conectando texto: Guía para el uso efectivo de elementos conectores en castellano*. Barcelona: Octaro.
- Montolio, Estrella. 2001. *Conectores de la lengua escrita: Contraargumentativos, consecutivos, aditivos y organizadores de la información*. Barcelona: Ariel.
- Noordman, Leo G.M. & Femke de Blijzer. 2000. "On the Processing of Causal Relations". *Topics in English Linguistics* 33.35–56. doi:10.1515/9783110219043.1.35
- Pander Maat, Henk & Liesbeth Degand. 2002. "Scaling Causal Relations and Connectives in Terms of Speaker Involvement". *Cognitive Linguistics* 12:3.211–245.
- Pander Maat, Henk & Ted Sanders. 2000. "Domains of Use or Subjectivity? The Distribution of Three Dutch Causal Connectives Explained". *Topics in English Linguistics* 33.57–82. doi:10.1515/9783110219043.1.57
- Portolés Lázaro, José. 1993. "La distinción entre los conectores y otros marcadores del discurso en español". *Verba: Anuario galego de filoloxia* 20.141–170.
- Portolés Lázaro, José. 2001. *Marcadores del discurso*. Barcelona: Ariel.
- Portolés Lázaro, José. 2005. *Pragmática para hispanistas*. Madrid: Síntesis.
- Real Academia Española. n.d. *Banco de datos CREA: Corpus de referencia del español actual*. Retrieved from <http://www.rae.es>.
- Real Academia Española. n.d. *Banco de datos CORDE: Corpus diacrónico del español*. Retrieved from <http://www.rae.es>.
- Reig Alamillo, Asela. 2011. "The Pragmatic Meaning of the Spanish Construction *lo que pasa es que*". *Journal of Pragmatics* 43:5.1435–1450. doi:10.1016/j.pragma.2010.10.015
- Rosy (username). n.d. "La gente de Rosy Runrún". Retrieved from [20minutos.es](http://20minutos.es).
- Sanders, Ted, José Sanders & Eve Sweetser. 2009. Causality, Cognition and Communication: A Mental Space Analysis of Subjectivity in Causal Connectives. *Causal Categories in Discourse and Cognition* ed. by Ted Sanders & Eve Sweetser, 19–59. Berlin: Mouton de Gruyter. doi:10.1515/9783110224429.19
- Sanders, Ted & Wilbert Spooren. 2010. "Discourse and Text Structure". *The Oxford Handbook of Cognitive Linguistics* ed. by Dirk Geeraerts & Hubert Cuyckens, 916–944. Oxford: Oxford University Press.
- Sweetser, Eve. 1990. *From Etymology to Pragmatics: Metaphorical and Cultural Aspects of Semantic Structure*. Cambridge & New York: Cambridge University Press. doi:10.1017/CBO9780511620904
- Travis, Catherine. 2005. *Discourse Markers in Colombian Spanish: A Study in Polysemy*. Berlin: Mouton de Gruyter.



## PART III

# Morphosyntactic variation and change





## *Yo no le conocí a mi abuela*

### The use of clitics *le*, *lo*, and *la* in Amazonian Colombian Spanish

Héctor Ramírez-Cruz

Universidad Nacional de Colombia / University of Pittsburgh

This paper treats the use of third-person pronominal objects *le*, *lo* and *la* in Amazonian Colombian Spanish. Pronoun selection was analyzed in relation to linguistic variables (case and grammatical gender of the co-referent), extra-linguistic variables concerning the speakers (age, gender and occupation) and a geographic variable (municipality). Results show that only linguistic variables have a significant effect on pronoun selection. In particular, we observe a significant interaction of case with grammatical gender, indicating a greater proportion of *leísmo* for feminine co-referents than for masculine co-referents. These findings contribute to understanding Spanish third-person pronominal objects in language contact, with additional insights about competition and second language learning strategies.

#### 1. Introduction

The Spanish third-person pronominal object system is composed of *le(s)*, *lo(s)* and *la(s)*. Unlike other grammatical persons, the third-person system is the only one that is able to distinguish case with different forms: canonically *lo* and *la* for accusative and *le* for dative. It is also the only one that distinguishes gender in accusative case and pluralizes with the suffix *-s*, as opposed to the first and second persons, which pluralize with different forms (e.g., *me* “I” / *nos* “we”) and have no gender distinction. This system originated in the rearrangement of pronouns and case markers in Late Latin, which produced the third-person anaphor *ille* (Wanner 1987: 75–76). This anaphor is the etymological basis for the third-person pronominal system in Romance languages, and it includes a diverse range of accusative and dative cases. The system is a source of variation both diachronically (Alfonso 1997) and synchronically (Fernández-Ordóñez 1999).

Using naturalistic data, this paper analyzes the use of the pronominal objects *le*, *lo* and *la* in Amazonian Colombian Spanish. The purpose of the paper is to study variation patterns on the selection of these pronouns in relation to linguistic variables (case and gender) and non-linguistic variables of the speaker (age, gender, occupation) and the geography (municipality). The corpus is composed of 68 spontaneous oral narrations of sixty-three Tikuna-Spanish fluent bilinguals recruited by the investigator using the snowball sampling method (Milroy & Gordon 2003: 32). Thus, generalizations of the results are constrained by the narrative style of the texts, as they favored the emergence of the pronouns due to frequent cross-references to the characters, spaces and objects of the narrations.

### 1.1 Context

This study was carried out in rural and urban indigenous settlements in Leticia and Puerto Nariño, Colombia as shown in Map 1 in the appendix, in 2005 and 2006. As the capital of the Department of Amazonas, Leticia had 42,877 inhabitants in 2005. Of them, 29,669 lived in the urban area and 13,208 in the rural communities (Asamblea Departamental, 2005: 11). Most of the non-indigenous inhabitants were highland Colombian immigrants established in the urban area, which also hosted local indigenous inhabitants. The rural area is dominated by indigenous people, who live in traditional communities. Puerto Nariño, on the other hand, is a smaller municipality located in the inner Amazon, 55 miles away from Leticia. At the time of fieldwork, this municipality had a smaller population of 7,190: 2,115 in the urban area and 5,075 in the rural communities (Asamblea Departamental 2005: 11). Tikunas were the dominant indigenous population in both municipalities with about 5,978 people (Gutiérrez et al. 2004: 234–243).

Located in Colombian borderlands, Leticia and Puerto Nariño have been socially integrated into the broader Amazon, which includes Brazilian, Ecuadorian, Colombian and Peruvian territories (Fenzl 2011: 30; Zárate 2011: 57). Leticia and Puerto Nariño have been especially linked to Peru and Ecuador via migration processes during the rubber boom (Domínguez 2004: 16; Marticorena 2004: 40). The patterns of migration indicate that the colonists entered through different rivers that connect the Andes and the Amazon. These migration processes may explain similarities between findings of this paper and uses of *le*, *lo* and *la* reported for the Andes and the Amazon in Peru. Iquitos was the main Peruvian city connected to Leticia and Puerto Nariño for extraction activities during the rubber boom in the Amazon (Ayarza 2004: 103; Arroyo 2004: 117; Flores 2004: 86; Riaño 2003: 20–21, 45–47, 81–82, 216–217). In addition, the national boundaries between different

countries on the Amazon were extensively debated until the 20th century (Zárate 2011: 71–72).

The establishment of different settlements in riverbank regions and the recruitment of indigenous people for the rubber industry may partially explain uses of *le*, *lo* and *la* found in the current speech of Tikunas. Nevertheless, demographic information presented by Alvar (1977: 381–382) shows that Peruvian immigrants surpassed the Colombian population in Leticia until 1946, when Colombian immigrants started to outnumber Peruvians. Census data from the 1990s indicates that about 3,659/22,866 (16%) and 146/5,490 (2.7%) of people in Leticia and Puerto Nariño, respectively, were Colombian immigrants. The data is complemented by census information from 2005 summarized in Table 1. These patterns of immigration comprise two main trends of immigrant population: Peruvian and Colombian (Aguilar 1992: 60; DANE 2005: 111–116; Sánchez 2012: 178).

**Table 1.** Colombian and Peruvian immigration in the twentieth century

Years	1935–1940	1941–1945	1946–1950	1990	2005
Colombian migrants	40	214	379	3,805	7,126
Peruvian migrants	76	226	205	286	423

## 1.2 Concepts and definitions

Regarding the distribution of third-person object pronouns in the Hispanic world, Fernández-Ordóñez (1999: 1320) names two types of zones: case distinction zones and zones which exhibit case confusion. The case distinction zones differentiate between accusative and dative cases with the pronominal objects *lo(s)* and *la(s)* for the former and *le(s)* for the latter. This system consistently distinguishes accusative and dative cases, distinguishes number agreement features and differentiates between masculine and feminine objects in accusative case. This is the etymological prescriptive system most diffused in standard varieties, including Colombian Spanish. Using data from the current study, in example (1) the speaker chooses the pronominal object *la* for a feminine determiner phrase *la muchacha* ‘the girl’, while *lo* is selected for a masculine singular phrase *el cuchillo* ‘the knife’ in example (2). On the other hand, the speaker chooses the pronominal object *le* when indexing a feminine phrase *la primera niña que usted tenga* ‘the first daughter that you have’ in example (3).<sup>1</sup>

1. These and further examples represent unedited utterances of study participants.

- (1) *A la muchachai la<sub>i</sub> pintan ahí, poco de gente ahí.*  
“They paint **the girl** there, a bunch of people there.”
- (2) *Ahora bótalo<sub>p</sub> el cuchillo<sub>i</sub>.*  
“Now throw **it** away, **the knife**.”
- (3) *Mi mamá me dijo “pero la primera niña que usted tenga<sub>p</sub> esa sí le<sub>i</sub> van., le<sub>i</sub> vamos a hacer [la pelazón]”*  
“My mother said to me ‘But we will do the haircut to **the first daughter that you have**.’”

For the zones with case confusion, either *le* is given totally or partially accusative functions or else *lo* or *la* are given dative functions. The first case is known as *leísmo*, which simplifies case distinctions with *le(s)* for both accusative and dative cases. As shown in example (4), the speaker chooses the pronominal object *le* as accusative case for a feminine phrase *mi abuela* “my grandma” and for a masculine pronoun *él* “he” in example (5). Therefore, the differentiation between accusative and dative cases does not persist as they merge in the unique form *le(s)*. *Leísmo* has been extensively documented both diachronically (Flores 2006; Alfonso, 1997) and synchronically (Fernández-Ordóñez 1999) in Spain and Latin America.

The second case corresponds to either *loísmo* or *laísmo*, which simplify case distinctions with the forms *lo(s)* or *la(s)*, respectively, for both accusative and dative cases. As shown in example (6), the speaker chooses the pronominal object *lo* as accusative for the feminine pronoun *ella* “she” and for a dative masculine phrase *su hermano* “his brother” in example (7). As a consequence, gender is simplified for the former and case is simplified for the latter. *Loísmo* and *laísmo* are less common cases which, however, have been documented in both Spain and Latin America, including the Amazon (Caravedo, 1997). Of those, *laísmo* is the rarest case.

- (4) *Yo no le<sub>i</sub> conocí a mi abuela<sub>i</sub>.*  
“I didn’t meet **my grandma**.”
- (5) *Él<sub>i</sub> tuvo un problema, por allá le<sub>i</sub> cogieron y está golpiao*  
“He got in trouble. Somebody caught **him** and he is injured.”
- (6) *El hermano de ella<sub>i</sub> se enamoró de ella<sub>i</sub> y lo<sub>i</sub> buscaba, buscaba, buscaba y nunca lo<sub>i</sub> encontraba*  
“Her brother fell in love with **her**, and he was looking and looking for **her** and he never found **her**.”
- (7) *Ese era castigo de su hermano<sub>p</sub> pues, o sea pa’ no hacerlo<sub>i</sub> mal nada, así lo castigó*  
“That was **his brother’s** punishment for him not to be bad, he punished him in that way.”

### 1.3 Mufwene's competition model

The arrangement and distribution of *le*, *lo* and *la* can be understood as a competition process between these units tied to a particular ecology. The competition is part of language change processes. Indeed, Mufwene (2001: 1–24) posits 'language evolution' as speciation processes via 'competition' of species in a given ecology, so that 'linguistic varieties' (languages and dialects) and 'variants' (specific linguistic features) are understood as species. In this paper, 'competition' is understood as an inherent process of linguistic diversification that can be maximized in language-contact situations. Specifically, competition is understood as a pool of linguistic features available for speakers to convey the same or similar functions. For example, the pronouns *le*, *lo* and *la* can compete for speakers' usage and result in different forms of 'case distinction', *léismo*, or *loísmo*. Similar competition processes are documented in this volume for *pa'/para* (see Fafulas et al., this volume), future tense forms (see Kyzar, this volume) and the complementizer *que/Ø* (see Ricelli, this volume).

'Ecology', on the other hand, is understood as a unique configuration of language and space. The ecology is usually shaped by unpredictable and non-replicable relationships between the language(s)/dialect(s) of a geographical space and the ethnohistorical and sociocultural conditions of the environment where they live. Under this framework, available linguistic units, such as *le*, *lo* and *la*, take part in contentious competitions of variable intensity as the index of particular ecologies. Dialect leveling (overlapping of different dialects) and koineization (mixing of dialects) are some of the processes that can take place in the competition among linguistic features (Mufwene 2001: 4–6).

The broader range of functions or meanings of each form is important when analyzing competing solutions since their motivations might be more closely related to semantic-cognitive functions in the speaker's mind than to morphosyntactic-grammatical factors in a linguistic system. Some of these factors/processes include animacy effects (Fernández-Ordoñez 1993: 5–6; Fernández-Ordoñez 1999: 1350–1355; Klee & Caravedo 2005: 15), cardinality and voice (Arnoux & Martínez 2001), continuity and definiteness (Fernández Ordoñez 1999: 1355–1360), transitivity (Flores 2006: 686–687), aspectual and pragmatic factors (Flores 2006: 688–706), the interaction of linguistic and external factors (Palacios 2007: 264–269) and social indexicalities (Mick & Palacios 2013). Although these factors may shape the results, an analysis of them is beyond the scope of this paper.

### 1.4 Languages in contact and the role of second language acquisition

The arrangement of the Spanish third-person object pronouns in bilingual contexts, such as the one examined in Leticia and Puerto Nariño, may be constrained by both the language contact setting (Thomason & Kaufman 1988: 65–109;

Thomason 2000: 70–71) and the strategies used by individuals when learning a second language (Andersen 1988). The former has been the dominant approach in accounting for clitic uses in Spanish-Indigenous language contacts in South and Central America. The latter has been supported by extensive research on L2 Spanish in instructional settings in environments in which English and other major L1s are spoken.

No previous research has been carried out on object pronouns in Amazonian Colombian Spanish. However, such studies have been conducted on neighboring areas, which are relevant for the current investigation as these areas share a geographical continuum with Leticia and Puerto Nariño, along with socio-historical relationships. For example, Caravedo (1997: 132–133, 140–141) documents the presence of both *leísmo* and *loísmo*, along with gender and number disagreement in the Peruvian Amazon. She analyzes a 360-minute oral corpus of 30-minute interviews with 12 literate and illiterate Spanish monolingual speakers.

Caravedo found 120 cases of *leísmo* among 249 pronominal constructions containing the pronoun *le(s)* and 7 cases of *loísmo* among 175 pronominal constructions containing the pronoun *lo(s)*; the remaining pronominal constructions belong to case distinction. These results show a stronger presence of *leísmo* with 48% (vs. 3% of *loísmo*); 85% of *leísmo* instances appeared in the illiterate group, while *loísmo* registers were evenly distributed between the literate and illiterate groups. This is the most relevant study on third-person pronoun objects from Amazonian Spanish. Importantly, the author explains the relationship between Andean and Amazonian Spanish through migration processes. Given such connections between these regions, some degree of dialect leveling among Peruvian and Colombian Spanish varieties may explain the arrangement of Spanish clitics in the communities surveyed.

Other authors have also documented *leísmo* and *loísmo* phenomena in the Peruvian Amazon. Barraza (1998: 26–27), for example, found *leísmo* as the dominant third-person pronominal use in Iquitos, a semi-urban city located in the Peruvian Amazon. She also found competing cases of *loísmo* under semantic and syntactic constraints, such as inanimate reference and preverbal position of the antecedent (Barraza 1998: 26–34, 41–47). However, other utterances did not select the *loísmo* variant under the same constraints. Similarly, Rojas (2004) found the dominance of *leísmo* as the most preferred third-person pronominal use in Lamas, a small Amazonian town in San Martín, Peru. This town is located in the highest part of the Amazon, known as the *selva alta*, which is a transitional region between the Amazon and the Andes.

Furthermore, a larger range of competing uses has been documented in the Andes. They include *leísmo*, *loísmo* and *laísmo*, as well as distinction and deletion of third-person pronominal objects in Quechua-Spanish contact settings (Escobar

2000: 41; Godenzzi 1996: 71–99; Martínez 1996: 139–177). If the Quechua-Spanish contact were the source of Peruvian *leísmo*, *loísmo* and *laísmo*, it is possible that such uses fossilized in the speech of bilinguals and then disseminated to monolinguals, such as those reported by Caravedo (1997) in the Peruvian Amazon or Tikuna communities, who are not Quechua speakers.

However, similar simplification processes of third-person pronominal objects have been reported in other regions of contact between Spanish and indigenous languages, such as Guaraní in Paraguay (Palacios 2000) and Tzutujil in Guatemala (García 2006). These findings raise the question of whether all cases of simplification of the third-person pronominal object system, extensively reported in different Spanish varieties of the Americas, may result from contact with local indigenous languages. In such a case, Spanish would tend towards such simpler systems and the simplification processes would be triggered or accelerated by the local indigenous languages, producing the same or similar predictable results in different geographical locations. Although the interpretation of these phenomena cannot be reduced to finding similarities between the grammatical inventories of the languages in contact, it is important to take this factor into account.

In the Amazon region surveyed, Spanish is an L2 among Tikuna communities. As the national language of Colombia, it serves official functions and acts as a lingua franca for communication with other ethnic groups. Tikuna, on the other hand, is the L1 of all participants and is the dominant indigenous language in the region examined. The third-person pronominal system of Tikuna is summarized in Table 2. This system distinguishes subjects with a different form for masculine and feminine singulars and a unique form for plurals. The system features agreement with verbal prefixes. According to Montes (2004a: 162–165), it assigns accusative case by affixation of the suffix *-û* to either a noun or a personal pronoun. The suffixes *-kà* and *-na* supply dative functions for benefactives and goals, respectively. The Tikuna system is illustrated in examples (8)–(9) taken from Montes (2004b: 125). In these examples, accusative (*-û*) and benefactive (*-kà*) affixes coexist in the same sentence and can appear in alternative order.

**Table 2.** Tikuna third-person pronominal system

3rd-person pronominal system		Verbal prefix	Subject pronoun	Object suffixes
Singular	Masculine	<i>ni-/na-</i>	<i>nüma</i>	<i>-û</i> (ACC)
	Feminine	<i>ngi-</i>	<i>ngímà</i>	<i>-kà</i> (BEN)
Plural		<i>ta-/ti-/to-</i>	<i>tüma</i>	<i>-na</i> (GOAL)



- (8) ĭt- ũ                      Gima-kà   chà-wa                      (Montes, 2004b: 125)  
 chontaduro-ACC Gima-BEN 1SG/PL-pick  
 “I pick chontaduro for Gima.”
- (9) Gima-kà   ĭt- ũ                      chà-wa                      (Montes, 2004b:125)  
 Gima-BEN chontaduro-ACC 1SG/PL-pick  
 “I pick chontaduro for Gima.”

Regarding language development, Montrul’s (2010) study of the use of clitics among Spanish heritage speakers and adult second language learners in the U.S. shows differences between adult learners of L2 Spanish, who had late exposure to Spanish in instructional contexts, presumably showing explicit knowledge of Spanish clitics, and Spanish heritage speakers, who had earlier exposure to Spanish in naturalistic contexts and displayed more intuitive knowledge of Spanish clitics (Montrul 2010: 198). These results may inform the Spanish clitic uses of Tikunas who had early exposure to Spanish without formal instruction on this particular structure.

Despite scholastic, theoretical and methodological differences between language contact studies (LCS), with a wider perspective of the social context, and second language acquisition (SLA) studies, with a narrower focus on the actual speakers’ processes, both approaches overlap in substantive matters (e.g., L1 transfer) and, therefore, they need to complement each other. For example, uses of Spanish clitics in Amazonian Spanish might be triggered either by a Tikuna L1 structure converging with a Spanish system or might be the result of a simplification process towards a dominant third-person pronoun in the input of Spanish L2 speakers. Both language transfer and simplification processes concern LCS and SLA approaches.

Furthermore, a transfer effect from L1 Tikuna favored by the social conditions of language contact and a cluster of SLA strategies might be acting together to produce a given result of clitic uses. For example, masculine is the default gender in Spanish (Franceschina 2005: 100), whereas feminine is the unmarked gender in Tikuna (Montes 2004b: 63). Therefore, while Tikuna case markers and Spanish clitics can converge via positive transfer from L1 (Siegel 2006: 25), gender values of Spanish clitics may be in conflict between the default masculine gender of Spanish and the generic feminine gender of Tikuna via simplification processes (Siegel 2006: 15). The competing variants of Spanish, which fuse gender in accusative clitics *lo* and *la*, and Tikuna, in which gender affixes are separated from case markers, are well understood by Mufwene’s (2001) competition model. Although elucidating the exact sources of clitic uses in Amazonian Spanish is beyond the scope of this paper, they may have multiple origins which are not mutually exclusive.

## 2. The present study

This chapter investigates pronoun selection of third-person pronominal objects: *le, lo* and *la*. It analyzes the dependence of this variable on linguistic (case and gender) and non-linguistic variables of the speaker and the geography. The interaction between variables is also analyzed. For the interpretation of results, historic processes of migration and language contact features are taken into account. The variables under analysis in the present study are summarized in Table 3. The study will give empirical data on the use of third-person pronouns from an undocumented region. It will add naturalistic evidence on the use of these pronouns in contexts of dialect and language contact with additional insights about competition. In light of this purpose and the background provided, the following research questions guided the present study:

- Is there any effect of linguistic variables on third-person pronominal object selection?
- Is there any effect of non-linguistic variables on pronominal object selection?
- What patterns of pronominal object selection emerge from the interaction of variables?

**Table 3.** Variables analyzed

Dependent variable	Independent variables	Non-linguistic independent variables	
Object pronouns	Linguistic variables	Speaker variables	Geographical variable
LE	Case	Age	Municipality
LO	Word gender	Gender	
LA		Occupation	

## 3. Methods

This paper uses ethnographic techniques for collection of data and combines them with quantitative methods for the analysis. The investigator collected and transcribed natural spoken data in a bilingual language-contact setting that included small rural indigenous villages and larger urban and semi-urban areas in Amazonas, Colombia. The present analysis involves a repeated measures unbalanced design with nested data, given that the size of the speech sample and the number of observations were not the same for all the subjects. Clitic variation in the third-person pronominal object system is examined using a mixed-effects model, further explained in the Results section.

3.1 Fieldwork

I conducted five weeks of fieldwork in rural and urban areas of the municipalities of Leticia and Puerto Nariño between December 2005 and January 2006, as shown on Map 1 in the appendix. I traveled to the urban towns of each municipality and six rural Spanish-Tikuna bilingual communities, located at the banks of the Amazon and Loretoyacu rivers. Most of the people in rural communities hold to traditional activities for survival such as fishing, farming, hunting and weaving, so that natural resources such as rivers are fundamental for daily activities (Rodríguez 2011). Those in the urban areas tend to develop more stationary lifestyles than those in the rural regions (Ochoa 2011).

During fieldwork, I recorded life stories, tales, open interviews about dreams, happy and sad events and personal anecdotes in Spanish. Subjects were prompted to narrate any story that they wanted to tell, be it traditional tales, their life stories or the history of the city/community they lived in. The corpus contains different modalities of direct and indirect discourse, past and present references and indicative and subjunctive moods, so that the narrators explored a variety of Spanish structures.

3.2 Population and sample

A total of 93 Tikuna-Spanish fluent bilingual subjects were interviewed and/or recorded in different situations. From those, a sample of 63 participants was selected, discarding speakers whose recordings were too short or had poor sound quality. The participants selected came from both Leticia (36/63) and Puerto Nariño (27/63). The majority of them (49/63) were located in rural areas, and 14/63 lived in the urban (Leticia city) or semi-urban (Puerto Nariño) towns. Table 4 summarizes the number of participants by community, whose locations can be found on Map 1 in the appendix.

Table 4. Subjects by community

Subjects by communities by municipality		Frequency	Percent (%)
Leticia	San José del Kilómetro 6	8	12.7
	Nazareth	13	20.6
	Arara	8	12.7
Puerto Nariño	Puerto Esperanza	6	9.5
	Nuevo Paraíso	6	9.5
	Santarén	8	12.7
Both Leticia & Puerto Nariño	Towns (urban)	14	22.2
	Total	63	100.0

The participants selected were adult Tikuna-Spanish bilinguals normally distributed at an average age of 44, ranging from 15 to 79 years of age. Fifty-two of them (83%) carry out traditional activities for survival such as fishing, hunting, farming, weaving and handicrafts, while eleven (17%) hold clerical positions (e.g., nurses, health promoters, administrative positions, teachers, students) or combine them with traditional activities.

### 3.3 Corpus

From the corpus of recordings carried out, a sample of 68 of these was selected; some study participants were represented with more than one recording in the sample. Only interviews containing tokens of pronominal third-person objects were included. The corpus totals 20 hours of natural speech recordings and 164,327 transcribed words. Given the size of the sample and of the corpus, the sample is arguably representative of Amazonian Colombian speech in narrative style. As shown in the following results, the number of tokens of third-person pronominal objects was also enough to draw conclusions about the use of these pronouns in the Colombian Amazon.

### 3.4 Data analysis

For the analysis of pronominal structures with *le, lo* and *la*, all tokens of these pronouns were located in the corpus. The nominal referents of those pronouns were identified in order to determine if they were feminine or masculine and accusative or dative. When doing this coding, the investigator did not assume any prescriptive position. Therefore, if the speaker said *un inyección* ‘a shot’ or *el inyección* ‘the shot’, that phrase counted as masculine even if *una inyección* and *la inyección* are feminine phrases in standard Spanish.<sup>2</sup> Thus, what took precedence was the gender assigned by the speaker and not a prescriptive norm. Given the very low counts of third-person plural pronouns, the number distinction was not taken into account.

Similarly, the case assigned to each pronoun depended on the semantic roles of the antecedent and their mapping to the pronominal structure. Although semantic roles are not entirely discrete (Dowty 1991: 572–576), accusative cases usually

---

2. In complex cases such as *la inyección doloroso*, I followed Franceschina (2005: 72, 84–87) assuming that the noun *inyección* is inherently [+fem]. The syntax of this phrase suggests that the feminine gender has reached the target *la*, with which the noun agrees, but not the target *doloroso*. So, *la inyección doloroso* would count as feminine. *El inyección doloroso*, which is another piece of the puzzle, should count as masculine because the intrinsic feature [+fem] of *inyección* is not visible in the targets.

map patient and theme roles that undergo a change of state, objects affected by an agent or cause or objects that are stationary with respect to the movement of another participant. Therefore, in a given structure such as example (4), the pronoun *le* counted as accusative, given that the coindexed referent *mi abuela* “my grandma” is represented as the object of another’s experience. Based on Campos (1999: 1519–1574), this would be an ‘(un)known object’ completing the meaning of the verb *conocer* “to meet” in a predication relationship. Similarly, in example (2), the pronoun *lo* counts as accusative given that the coindexed referent *el cuchillo* “the knife” is represented as an object undergoing a change of location caused by the action of another participant.

Likewise, datives usually map goals, beneficiaries and experiencers (Haspelmath 2003: 213). For instance, in example (3), *le* was taken as dative given that the coindexed phrase *la primera niña que usted tenga* “the first daughter that you have” is represented as someone that receives the benefit or prejudice of an action (the haircut) performed by another participant. Once these analytical decisions were made, data were coded and submitted to a mixed-effects model for statistical analysis, as shown in the next section.

## 4. Results

The data were numerically inspected and analyzed using a mixed-effects model as implemented in the Rbrul program (Johnson 2009). This method appropriately handles repeated measures with nested data in non-normal distributions. The model is appropriate for these data as they come from spontaneous oral narrations of variable duration per speaker. Indeed, the number of tokens of pronominal objects per narration fluctuates from low to high frequencies, since a few speakers did not use these pronouns and other speakers used them to a large extent. These are expected behaviors of naturalistic data, which are accurately addressed by the model. For more on the advantages of mixed-effects models in linguistic research, see Fafulas et al. (this volume).

### 4.1 Statistical analysis

As shown in Table 5, a total of 2,888 tokens of third-person pronominal objects were found in the corpus. Most of them were instances of *le* with 2,192 occurrences (76%,  $\bar{X}_1 = 34.8$ ). *Lo* and *la*, on the other hand, comprised the remaining data, with the lowest proportion consisting of *la* (77 tokens, 3%,  $\bar{X}_2 = 1.2$ ) and a much higher proportion of *lo* (619 tokens, 21%,  $\bar{X}_3 = 9.8$ ). These data show that *le* was the most frequent choice, while *la* seems to have virtually vanished.

For the mixed-effects model, Speaker was set as a random factor and Pronoun was set as a dependent categorical variable with two levels: (1) *le* and (2) *lo* / *la*; thus *lo* and *la* were merged as a single level. This merge is justified by the low tokens of *la* and the fact that *lo* and *la* canonically distinguish gender for accusative case. Two independent linguistic variables were analyzed: (1) Case with 2 levels, accusative and dative, and (2) Word Gender with two levels, masculine and feminine gender of a noun coindexed with the pronoun. The non-linguistic independent variables were (3) Age, (4) Gender, (5) Occupation of the speaker and the (6) Municipality where the speaker lived. The interaction between variables was also analyzed. Figure 1 displays the variables included in the model.

**Table 5.** Tokens per pronoun

	LE_Total	LO_Total	LA_Total	TOTAL
Mean	34.79	9.83	1.22	
Sum	2192	619	77	2888
%	76%	21%	3%	100%

The output from Rbrul is summarized in Tables 6, 7 and 8. Table 6 contains the output for the independent variables. Reading from left to right, the second column includes the number of tokens (pronoun tokens) in each level of each factor. The third column corresponds to the coefficients of comparison, with positive coefficients favoring *le*. The forth column indicates the proportion of application of each level of each variable to *le*. The next column corresponds to the weights of each level from 0 to 1, in which numbers above .5 favored *le*. The larger the range of these numbers, the more likely the factor is to be significant. The last two columns include *p* and *r*<sup>2</sup> values as measures of significance and proportion of variation accounted by the factor, respectively. Significant values were flagged with an asterisk (\*). Table 7 displays selection of the pronoun *le* vs. *lo/la* for both feminine and masculine gender in accusative and dative case. Table 8, on the other hand, summarizes the best cumulative step-up model given by the program, including all significant factors, a significant interaction between case and word gender, degrees of freedom and proportion of variability accounted by the model.

As shown in Table 6, the non-linguistic variables were not significantly related to the pronouns selected: gender (*p* = .15), age (*p* = .17), occupation (*p* = .33) and municipality (*p* = .59). Only the linguistic variables gave significant values. Thus, case (*p* = 2.57e-201) and word gender (*p* = .00187) are significantly related to the pronouns chosen, as dative case and feminine gender of the coindexed nouns gave more weight to *le* than their counterparts, accusative and masculine.

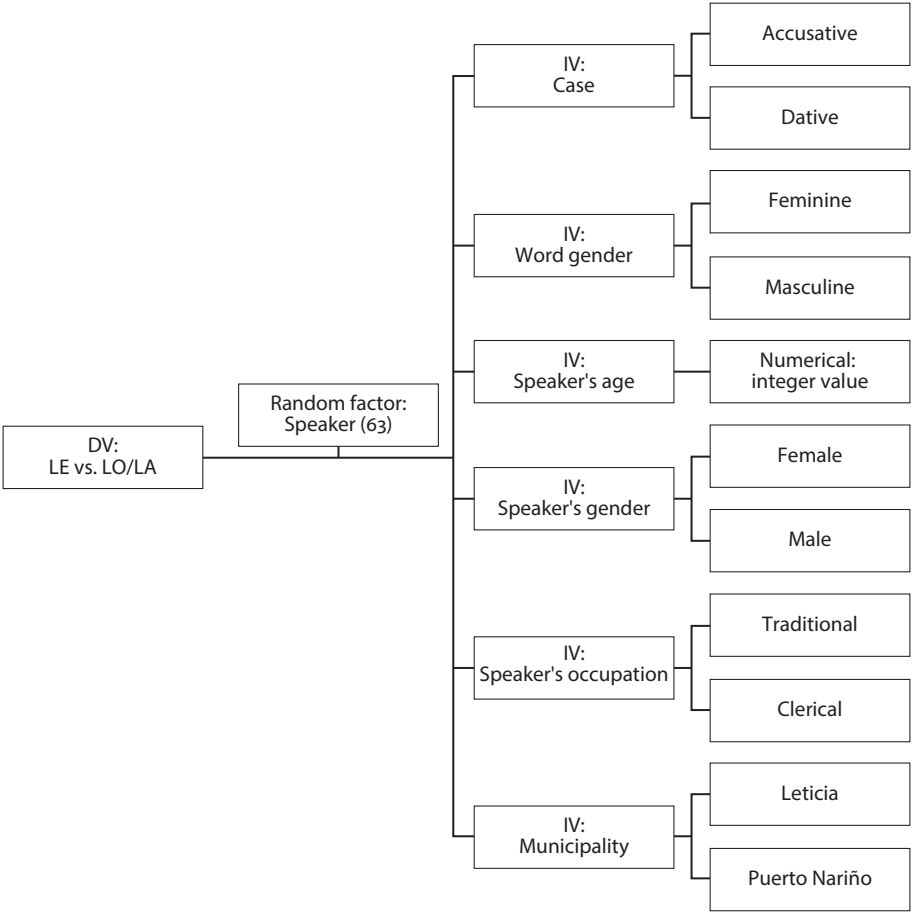


Figure 1. Variables studied

Table 6. Fixed independent factors on pronoun selection

Factors	Tokens (N)	Coefficients (Log Odds)	Proportion of application value	<i>p</i> -value *sig < .05	Weight	<i>r</i> <sup>2</sup>
Case				2.57e-201*		.417
Dative	1408	2.111	.973		.892	
Accusative	1480	−2.111	.555		.108	
Word gender				.00187*		.01
Feminine	1108	.053	.799		.513	
Masculine	1780	−.053	.734		.487	

Table 6. (Continued)

Factors	Tokens (N)	Coefficients (Log Odds)	Proportion of application value	Weight	<i>p</i> -value *sig < .05	<i>r</i> <sup>2</sup>
<b>Speaker's gender</b>					.153	.015
Female	1318	.269	.862	.567		
Male	1570	-.269	.673	.433		
<b>Speaker's age</b>	2888	.016	continuous variable (integers)		.172	.011
<b>Speaker's occupation</b>					.335	.009
Traditional	1831	.22	.777	.555		
Clerical	1057	-.22	.728	.445		
<b>Municipality</b>					.591	.002
Puerto Nariño	1226	.104	.822	.526		
Leticia	1662	-.104	.712	.474		

*LE* was set as the application value for each factor: *le* / (*le* + (*lo* & *la*))

The significance of case is explained by a categorical relationship between the pronoun *le* and dative case, which is the canonical function of this pronoun. The significance of gender, in which feminine gives significantly more weight to *le* than masculine, suggests a different distribution of the third-person pronouns by the co-referent word gender across case. Indeed, Table 7 shows a significant interaction between word gender and accusative case but not with dative case. These data indicate no significant variation for dative case, in which the speaker mostly selects *le* regardless of the co-referent word gender, as opposite to accusative case in which there was some variation between *le* and *lo/la* giving room for *leísmo*, that is, selection of *le* for accusative case.

Table 7. Selection of third-person pronoun in accusative and dative case

#### Selection of third-person pronoun LE vs LO/LA

*LE* was set as the application value for each factor: [*le* / (*le* + (*lo* & *la*))]

	Accusative			Dative		
Input	.523			.996		
Total N	1480			1408		
<b>Word gender</b>	<i>p</i> = .00019*			<i>p</i> = .249		
	N	FW	Proportion of application value (% LE)	N	FW	Proportion of application value (% LE)
Feminine	571	.576*	63.9	537	.439	96.8
Masculine	909	.424*	50.3	871	.561	97.6



Table 8. Best step-up mixed-effects model

Deviance	df	r <sup>2</sup> fixed	r <sup>2</sup> random	r <sup>2</sup> total
1566.617	5	.419	.292	.711

Speaker [random] + Case ( $p = 2.57e-201$ ) + Word.Gender ( $p = .00187$ ) + Case : Word.Gender ( $p = .0104$ )  
This model is significantly better with than without the interaction  
Case: Word.Gender,  $\chi^2 = 7.28$ ,  $df = 1$ ,  $p = .00695$

4.2 Pronoun selection among the levels of case across word gender

Figure 2 shows the patterns of pronoun selection according to case across word gender. Numbers above .5 indicate greater weight for pronoun *le*, and numbers below .5 indicate lower weight for *le* and, therefore, greater weight for selection of pronouns *lo* or *la*. The perimeter width depicts sample size per cell scaled at .045. The numbers of tokens per pronoun are displayed in the adjacent table.

As depicted in Figure 2, dative case gave greater weight for selection of *le*, and word gender made no difference in this particular trend. Thus, *le* was the most frequent pronoun for both masculine and feminine dative. This is the expected behavior of Spanish third-person object pronouns, in which the Colombian Amazonian speech examined seems to follow the general patterns of standard Spanish described by Fernández-Ordóñez (1999: 1386–1390). Example (10) illustrates *le* in the dative case for a masculine coindexed phrase *el mojoyoy*, an earthworm that is part of Tikuna cuisine.

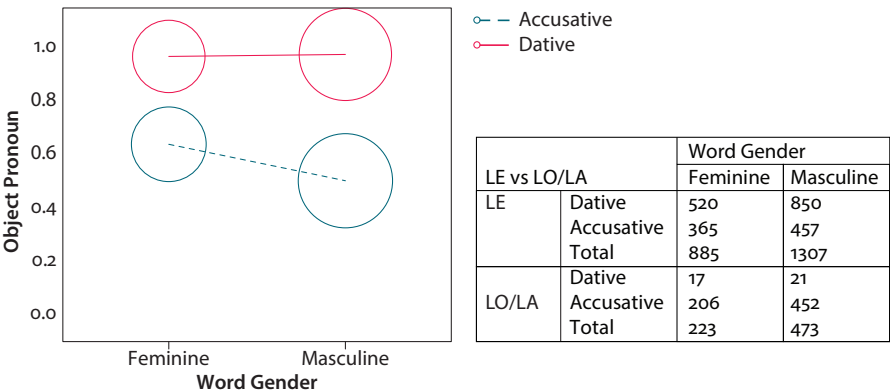


Figure 2. Pronoun selection by case across word gender

- (10) LE masculine dative:  
*Con el mojoyoyi ... mi mamá le exprimía la mantequita, el aceitico, de ese gusanito*  
“With mojoyoy ... my mom would wring the butter from it, the oil, from that earthworm.”

Similarly, example (11) illustrates selection of *le* in dative case for a feminine coindexed pronoun *ella* “she”, which is also coindexed to the feminine phrase *una muchacha* “the girl”. In the example, the narrator represents the voice of a young male character, as well as the voice of the character’s mother, who advises the male character. The example is particularly illustrative of the case distinction system, as the speaker chose *le* with a communicative verb *hablar* “to speak” but *la* with the monotransitive verb *llamar* “to call”. Both structures are adjacent to each other and relate to the same referent *una muchacha* “a girl”, suggesting that the speaker clearly assigns a different function to each pronoun.

(11) LE feminine dative:

FAC\_UPN37: “Encontré en el camino un rastro de *una muchacha*.” Su mamá dijo: “hijo tienes que ir y si ves otra vez el rastro, rastro de *ella*, las huellas de *ella*, entonces usted, háblele, llámela, a ver si se aparece.”

“I found some footprints of a girl in the road. Her mom said, ‘My son, you must go and if you see the prints again, her footprints, the footprints of her, then speak to her, call out to her, and see if she appears’”

On the other hand, Figure 2 points out where the significant difference of case across gender is, as that difference only concerns accusative case. Both genders are close to a positive weight for accusative *le*, but accusative feminine cases gave a greater weight for *le* than accusative masculine cases did. Although the pattern seems counterintuitive, it is explained by the very low tokens of pronoun *la* (3% of the corpus) as contrasted with *lo*, the overall generic use of *le* and possibly the unmarked status of feminine gender in Tikuna. This suggests that *le* was mostly selected for feminine coindexed nouns, regardless of case. Accusative masculine coindexed nouns, on the other hand, gave lower weight for *le*, as the pronoun *lo* seems still available for speakers with a proportion of 21% over the whole corpus. In fact, the perimeter width of accusative masculine in Figure 2 confirms a greater proportion of masculine referents as compared to feminine referents. The weighted mean close to .5 indicates that the likelihood of selecting either *le* (457/909) or *lo* (452/909) are almost equal, suggesting some competition between these pronouns in accusative case.

Example (12) illustrates an accusative feminine case with *le*. The speaker is describing the process of growing and preparing *fariña* “manioc”. She selected the pronoun *le* multiple times for the same referent with different transitive verbs: *le tenemos* “we have it”, *le embolsamos* “we put it inside plastic bags” and *le enterramos* “we bury it”. A similar example in (13) illustrates selection of *lo* for a masculine coindexed phrase *el maíz* “corn” in accusative case. In the example, the speaker is describing the process of growing corn and selects *lo* to index this phrase.

- (12) LE feminine accusative:

GAP\_UPN39: *la, la mata de la fariña<sub>i</sub>, nosotros le<sub>i</sub> tenemos como almacenado, le<sub>p</sub> le<sub>i</sub> embolsamos con la bolsa de plástico y le<sub>i</sub> enterramos*  
 “The, the plant, of manioc, we have it stored, we pack it, pack it in a plastic bag and we bury it.”

- (13) LO masculine accusative:

FAC\_UPN37: *el maíz<sub>p</sub>, él lo<sub>i</sub> hacía y lo<sub>i</sub> sembraba*  
 “Corn, he made it and grew it.”

In general, accusative case seems unstable for pronoun selection. This instability is illustrated in examples (14)–(16). In example (14), the speaker chose pronoun *la*, as expected, for a feminine coindexed phrase *una cartilla* “a booklet” in accusative case. However, *lo* and *le* are selected for accusative feminine phrases *la lengua* “the language” and *la muchacha* “the girl” in examples (15) and (16), respectively. The speaker’s hesitation regarding the pronoun in example (16) also suggests this instability.

- (14) LA feminine accusative:

WMD\_UNL31: *es una cartilla pequeña<sub>i</sub>, que la<sub>i</sub> hice yo, se perdió*  
 “It’s a small booklet that I made, it got lost.”

- (15) LO feminine accusative:

LRP\_ARLA: *esa es la primera lengua<sub>i</sub> [lengua géral] ... los demás indígenas ... por allá en, en Contreras, eso sí lo<sub>i</sub> hablan y en el Brasil ... esa lengua<sub>i</sub> viene del Brasil ... de, de, de, de Belén, para abajo. Esa gente sí lo<sub>i</sub> hablan.*  
 “That was the original language [general language] ... other indigenous people ... over there in, in Contreras, they do speak it and in Brazil ... that language comes from Brazil ... from, from, from, from Belén, down there. Those people do speak it.”

- (16) LE/LA feminine accusative:

DPA\_ARL7: *Entonces de ahí hacemos la masato y buscar la, una cosa, yanchama para hacer los juegos y le hacemos la masato y tomamos ahí y le<sub>p</sub> y la<sub>i</sub> tomamos pa’ pelar la, a la muchacha<sub>i</sub>.*  
 “So, that is how we prepare the masato [yucca drink] and we look for, one thing, yanchama [tree bark] in order to play and we prepare masato for her and we drink over there, and we take her, her, we take her to crop her hair, the girl’s hair.”

Isolated cases of *loísmo* also contributed to the competition of third-person pronouns. There were 38 tokens of dative *lo/la*. Of those, 16 tokens of dative *lo* were found in a single community: Santarén, the farthest indigenous community I visited in the municipality of Puerto Nariño, about 25 miles from the town of Puerto

Nariño and 80 miles from the Leticia city (see Map 1). Those were considered geographically isolated instances of *loísmo* (use of *lo* for dative case) that make up only 2.42% of all uses of *lo* but 16.66% of dative cases in Santarén and 42% of the 38 tokens of dative *lo/la*. The remaining tokens of dative *lo/la* are spread elsewhere. Example (17) illustrates the exceptional tokens of *loísmo* from Santarén. The speaker coindexes the pronoun *lo* with the feminine determiner phrase *las personas que salían del pescado* using the verb *poner* “to put”, which implies the imputation of the direct object *los clanes* “the clans” to a recipient: the people who received the clan names.

- (17) LO feminine dative (Santarén community, Puerto Nariño municipality):  
 ASA\_SRPN93: *Yoi lo<sub>i</sub> puso los clanes a las personas que salían del, del*  
*pescado<sub>i</sub>*  
 “Yoi gave **them** clan names to **people who emerged from fish**.”

Overall, this section has illustrated that only accusative case made a significant difference in pronoun selection. In spite of exceptional cases of dative *lo*, dative case made no significant difference in pronoun selection, as speakers categorically selected *le* for dative case. Similarly, word gender had no effect on pronoun selection in dative case, in which speakers consistently selected *le* regardless of gender. However, there was more variation in pronoun selection for accusative case, which triggered the significant difference found by the statistical tests, with an important proportion of *leísmo*: 37.5% of all uses of *le* (822/2192). Furthermore, the significant interaction between case and word gender shows accusative feminine cases giving greater weight for *le* than accusative masculine cases. In short, there was less variation in dative case because gender distinction was not relevant in this case, whereas that distinction did appear to be relevant for accusative case.

## 5. Discussion

These findings provide naturalistic evidence of the use of third-person pronominal objects in two municipalities of the Colombian Amazon, Leticia and Puerto Nariño. The patterns of case difference have indicated the greatest variation in accusative case, in which a significant proportion of *leísmo* has been found especially for feminine co-referents (63.9% out of 571 tokens). There was a lower but still important proportion of masculine co-referents of *le* (50.3% out of 909 tokens). A few cases of *loísmo* have been also found, most of them in an isolated indigenous community.

The migration patterns suggest that *leísmo* and *loísmo* are dialectal features of Andean Spanish brought to the Amazon by Peruvian *mestizos* during the rubber boom in the late 19th and early 20th centuries. These features arrived in Leticia

and Puerto Nariño from Iquitos, Peru, which was the main enclave of the rubber industry in the Amazon. Given that the specific extractive activities developed in Leticia and Puerto Nariño were controlled from Iquitos (Domínguez 2004:16; Marticorena 2004: 40; Arroyo 2004: 117; Riaño 2003: 20–21, 45–47, 81–82, 216–217), this is the main possible center of diffusion of Andean Spanish features into the Colombian Amazon. These facts explain the similarities between findings of this paper and those reported for nearby areas in the Peruvian Amazon (Caravedo 1997; Barraza 1998; Rojas 2004) and suggest dialectal contiguity across these areas. These findings contribute to the cumulative knowledge on Spanish third-person pronominal simplification in language contact areas from Latin America (Escobar 2000; Palacios 2000, 2007; García 2006; Godenzzi 1996).

Likewise, the strengthening presence of Colombia in Leticia, which earlier belonged to Peru, may explain the lower presence of *leísmo* (28.4% of all data) and isolated cases of *loísmo* as compared to the case distinction system (70.2% of all data), holding pronoun *le* for dative case and *lo/la* for accusative case. Indeed, a quick review of spoken data of Colombians living in Leticia in the 1970s (Alvar 1977: 307–349) reveals 32/37 (86.5%) cases of case distinction, 5/37 (13.5%) cases of *leísmo* and 0/37 cases of *loísmo*. Case distinction was the most plausible pattern brought in by the Colombian migrants. However, the depiction of the diffusion patterns of these norms – case distinction, *leísmo* and *loísmo* – would need further examination.

Results have been interpreted in light of Mufwene's (2001) competition model. There were three competing variants: *le*, *lo*, *la*. Of them, *la* seems virtually eliminated with the lowest percentage (3%), *lo* stood in second place (21%) and *le* was the dominant form (76%). *Le* comprises the majority (97.3%) of all dative cases – which is its canonical function, a significant proportion (55.5%) of all accusative cases (*leísmo*) and a significant part (63.92%) of feminine accusatives (*leísmo*). With respect to the forms *lo/la*, these pronouns comprise 696 tokens (24% of the data). A great majority (94.5%) of these tokens are accusative, leading case distinction, whereas a minor proportion (5.4%) are dative. Dative functions of *lo/la* correspond to isolated cases of *loísmo*, mainly from Santarén in Puerto Nariño. All in all, the outputs of this competition took the form of case distinction, *leísmo* (which brings simplification of case and/or word gender) and isolated cases of *loísmo*. The coexistence of these outputs is evidence of both dialect leveling and koineization in Amazonian Colombian Spanish. Kyzar (this volume) also indicates some possible dialect leveling between Spanish varieties in Louisiana. In Mufwene's (2001: 4–6) model, all these facts represent restructuring processes that take place in language evolution/change.

Simplification of case and gender may correlate with SLA strategies or with some influence from the main language of contact: Tikuna. As an SLA strategy, it is reasonable to posit simplification processes towards a generic Spanish form *le*, perhaps fossilized in bilingual speakers. Furthermore, the pronoun *le* in accusative

function comprises a larger proportion of all feminine tokens (63.9% out of 571 tokens) than the proportion comprised for all masculine tokens (50.3% out of 909 tokens). Given the larger proportion of accusative feminine cases, it is also reasonable to state some possible influence from the unmarked feminine gender in Tikuna, in which the object suffixes (Montes 2004b: 125) do not fuse gender information. However, the simplification processes observed are not direct evidence of influence from Tikuna, given the geographic and socio-historical contiguity with the nearby Peruvian Amazon, where coincident phenomena have been reported. Ultimately, some possible influence from Tikuna is not incompatible with these results and it can be one of multiple factors contributing to the documented simplification processes.

## 6. Conclusion

This paper has analyzed the use of the third-person object pronouns in Amazonian Spanish. The results were based on oral narrations representative of the speech from rural and urban indigenous settlements in Leticia and Puerto Nariño, Colombia. The three research questions posed at the outset were answered as follows: (1) there was an effect of case and grammatical gender of the co-referent on pronominal object selection; (2) there was no effect of non-linguistic variables on pronoun selection; and, finally, (3) patterns of pronoun selection showed variation for accusative case across the co-referent grammatical gender. Thus, *léismo* was dominant for accusative feminine co-referents, whereas the pronouns *le* and *lo/la* alternate more evenly for accusative masculine co-referents. These results are the byproduct of migration processes from Peru and Colombia and can be analyzed from the perspective of Mufwene's (2001) competition model, in which SLA strategies and some possible influence from Tikuna might play a role.

## Acknowledgements

The fieldwork for this research was funded by the Universidad Nacional de Colombia through the project *Bilingualism and Language Contact: Case Studies on Amazonian Languages*. I am very grateful to the University; the project's director, Maria Emilia Montes; and the Department of Linguistics at the University of Pittsburgh for the funding provided to present these findings at the XXIX Biennial Louisiana Conference on Hispanic Languages and Literatures in 2014. I especially thank Daniel Ezra Johnson for his assistance with the statistical tools and illustrations. Thanks to Hillary Sheeps and Lauren Posey for carefully proofreading this paper. All remaining errors and omissions are my responsibility.

## References

- Aguilar, Nelly. 1992. *Poblamiento de la Amazonia colombiana*. Santiago de Chile: Centro Latinoamericano de Demografía.
- Alfonso, Milagros. 1997. "Tres motivaciones para un cambio: El caso de los clíticos en las causativas con infinitivo en el español medieval". *Cambios diacrónicos en el español* ed. by Concepción Company, 11–31. Mexico City: Universidad Nacional Autónoma de México.
- Alvar, Manuel. 1977. *Leticia: estudios lingüísticos sobre la Amazonia colombiana*. Bogotá: Instituto Caro y Cuervo.
- Andersen, Roger W. 1988. "Models, Processes, Principles, and Strategies: Second Language Acquisition in and out of the Classroom". *Issues and Developments in English and Applied Linguistics (IDEAL)* 3.111–138.
- Arnoux, Elvira & Angelita Martínez. 2001. "Del oído al ojo: La variación en la producción escrita". *Estudios de Lingüística Española* 13. Retrieved from <http://elies.rediris.es/elies13/arnoux.htm>.
- Arroyo, José M. 2004. "Notas sobre el habla de Loreto". *Kanatari* 1050.107–122.
- Asamblea Departamental. 2005. *Plan de desarrollo del Departamento del Amazonas*. Leticia, Col.: Gobernación del Departamento de Amazonas.
- Ayarza, Armando. 2004. "La literatura en la Amazonia peruana". *Kanatari* 1050.101–103.
- Barraza, Yris. 1998. *Apuntes sobre gramática del castellano de Iquitos*. Lima: Universidad Nacional Mayor de San Marcos.
- Campos, Héctor. 1999. "Transitividad e intransitividad". *Gramática descriptiva de la lengua española* ed. by Ignacio Bosque & Violeta Demonte, 1519–1574. Madrid: Espasa Calpe.
- Caravedo, Rocío. 1997. "Los pronombres objeto en un corpus del español amazónico peruano". *Anuario de Letras* 35.131–155.
- Departamento Administrativo Nacional de Estadística (DANE). 2005. *Censo General 2005. Nivel Nacional*. Bogotá: DANE.
- Domínguez, Camilo. 2004. "Prefacio". *Perfiles urbanos en la Amazonia colombiana* ed. by Franz Gutiérrez, Luis E. Acosta & Carlos A. Salazar, 15–16. Bogotá: Guadalupe.
- Dowty, David. 1991. "Thematic Proto-Roles and Argument Selection". *Language* 67:3.547–619. doi:10.1353/lan.1991.0021
- Escobar, Anna M. 2000. *Contacto social y lingüístico*. Lima: Pontificia Universidad Católica del Perú.
- Fenzl, Norbert. 2011. "Integración de la Amazonia: Desafíos, obstáculos y perspectivas". *Amazonia colombiana: Imaginarios y realidades* ed. by Juan A. Echeverri & Catalina Pérez, 27–35. Leticia, Col.: Universidad Nacional de Colombia.
- Fernández-Ordóñez, Inés. 1993. "Leísmo, laísmo y loísmo: estado de la cuestión". *Los pronombres átonos* ed. by Olga Fernández, 63–96. Madrid: Taurus Universitaria.
- Fernández-Ordóñez, Inés. 1999. "Leísmo, laísmo y loísmo". *Gramática descriptiva de la lengua española* ed. by Ignacio Bosque and Violeta Demonte, 1317–1397. Madrid: Espasa Calpe.
- Flores, Guillermo. 2004. "El lenguaje global en los medios y la tradición local". *Kanatari* 1050.85–88.
- Flores, Marcela. 2006. "Leísmo, laísmo, loísmo". *Sintaxis histórica de la lengua española* ed. by Concepción Company, vol. 1, 671–749. Mexico City: Fondo de Cultura Económica & Universidad Nacional Autónoma de México.

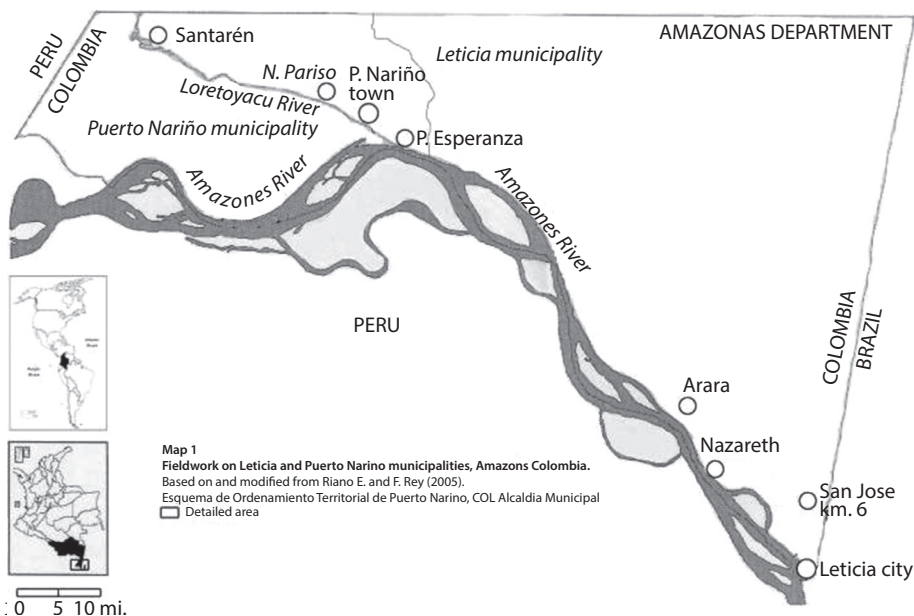


- Franceschina, Florencia. 2005. *Fossilized Second Language Grammars: The Acquisition of Grammatical Gender*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/lald.38
- García, Ana I. 2006. "Contacto de lenguas en Guatemala: Cambios en el sistema pronominal átono del español por contacto con la lengua maya Tzutujil". *Tópicos del Seminario* 15.11–71.
- Godenzzi, Juan C. 1996. "Transferencias lingüísticas entre el quechua y el español". *Signo y Seña* 6.71–99.
- Gutiérrez, Franz, Luis E. Acosta & Carlos A. Salazar. 2004. *Perfiles urbanos en la Amazonia colombiana*. Bogotá: Instituto Amazónico de Investigaciones Científicas.
- Haspelmath, Martin. 2003. "The Geometry of Grammatical Meaning: Semantic Maps and Cross-Linguistic Comparison". *The New Psychology of Language: Cognitive and Functional Approaches to Language Structure* ed. by Michael Tomasello, vol. 2, 211–242. London: Erlbaum Associates.
- Johnson, Daniel E. 2009. "Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis". *Language and Linguistic Compass*, 3:1.359–383. doi:10.1111/j.1749-818X.2008.00108.x
- Klee, Carol A. & Rocío Caravedo. 2005. "Contact-Induced Language Change in Lima, Perú: The Case of Clitic Pronouns". *Selected Proceedings of the 7th Hispanic Linguistics Symposium* ed. by David Eddington, 12–21. Somerville, Mass.: Cascadilla.
- Marticorena, Manuel. 2004. "La novela amazónica y sus variantes". *Kanatari* 1050.38–43.
- Martínez, Angelita. 1996. "Lenguas y culturas en contacto: Uso de los clíticos *lo-la-le*". *Signo y Seña* 6.139–177.
- Mick, Carola & Azucena Palacios. 2013. "Mantenimiento o sustitución de rasgos lingüísticos indexados socialmente: Migrantes de zonas andinas en Lima". *Lexis* 37:2.341–380.
- Milroy, Lesley & Matthew Gordon. 2003. *Sociolinguistics: Method and Interpretation*. Malden, Mass.: Blackwell. doi:10.1002/9780470758359
- Montes, María E. 2004a. "Lengua ticuna: Resultados de fonología y morfosintaxis". *Forma y Función* 17.145–178.
- Montes, María E. 2004b. *Morfoxintaxis de la lengua tikuna (Amazonia colombiana)*. Bogotá: Universidad de los Andes.
- Montrul, Silvina. 2010. "How Similar are Adult Second Language Learners and Spanish Heritage Speakers? Spanish Clitics and Word Order". *Applied Psycholinguistics* 31.167–207. doi:10.1017/S014271640999021X
- Mufwene, Salikoko S. 2001. *The Ecology of Language Evolution*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511612862
- Ochoa, Germán I. 2011. "Ciudades, ambiente y diversidades urbanas en la Amazonia". *Amazonia colombiana: Imaginarios y realidades* ed. by Juan A. Echeverri & Catalina Pérez, 391–403. Leticia, Col.: Universidad Nacional de Colombia.
- Palacios, Azucena. 2000. "El sistema pronominal del español paraguayo". *Teoría y práctica del contacto: El español de América en el candelero* ed. by Julio Calvo, 122–143. Frankfurt & Madrid: Vervuert/Iberoamericana.
- Palacios, Azucena. 2007. "¿Son compatibles los cambios inducidos por contacto y las tendencias internas del sistema?" *La Rumania en interacción: Entre historia, contacto y política* ed. by Martina Schrader-Kniffki & Laura Morgenthaler, 259–279. Frankfurt & Madrid: Vervuert/Iberoamericana.
- Riaño, Elizabeth. 2003. *Organizando su espacio, construyendo su territorio*. Leticia, Col.: Universidad Nacional de Colombia.



- Rodríguez, Carlos. 2011. "Visiones y manejo local del mundo del agua en la Amazonia colombiana". *Amazonia colombiana: Imaginarios y realidades* ed. by Juan A. Echeverri & Catalina Pérez, 171–181. Leticia, Col.: Universidad Nacional de Colombia.
- Rojas, Pilar. 2004. *Caracterización general del español hablado por los indígenas kechwas del barrio Wayku en la Provincia de Lamas, departamento de San Martín, Perú*. Leticia, Col.: Universidad Nacional de Colombia.
- Sánchez, Aníbal. 2012. *Perfil migratorio del Perú 2012*. Lima: Organización Internacional para las Migraciones.
- Siegel, Jeff. 2006. "Links between SLA and Creole Studies". *L2 Acquisition and Creole Genesis* ed. by Claire Lefebvre, Lydia White & Christine Jourdan, 15–46. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/lald.42.03sie
- Thomason, Sarah G. 2000. *Language Contact: An Introduction*. Washington: Georgetown University Press.
- Thomason, Sarah G. & Terrence Kaufman. 1988. *Language Contact, Creolization, and Genetic Linguistics*. Los Angeles: University of California.
- Wanner, Dieter. 1987. *The Development of Romance Clitic Pronouns from Latin to Old Romance*. Berlin: Walter de Gruyter. doi:10.1515/9783110893069
- Zárate, Carlos G. 2011. "Amazonia: La historia desde la frontera". *Amazonia colombiana: Imaginarios y realidades* ed. by Juan A. Echeverri & Catalina Pérez, 55–76. Leticia, Col.: Universidad Nacional de Colombia.

## Appendix



# Attitudes toward morphosyntactic variation in the Spanish of Valencian speakers

Rosa María Piqueres Gilabert & Matthew Fuss  
Indiana University

Blas Arroyo (1991, 1993, 1999, 2011) and González Martínez & Blas Arroyo (2011) identified degrees of variation in the morphosyntax of the Spanish of Valencia. Blas Arroyo (1993) observed variation in the use of locative prepositions *a/en* “at/in”, partitive *de* “of” and expletive *que* “that” in interrogative sentences. This paper describes how these variables are evaluated, using the matched-guise technique (Lambert et al. 1960). Four Valencian guise speakers from Castellón pronounced 12 sentences containing these three variables. A matched-guise test was completed to determine the attitudes of 31 Valencian-Spanish bilinguals toward the variation in these patterns. Results reveal a significant difference in the socioeconomic status evaluation for locative *a/en*, while all non-standard variants share positive evaluations for the personal attractiveness sphere. There is no gender or quality voice effect, which reinforces the validity of this instrument. Findings are discussed to understand the meaning of morphosyntactic variation in Castellón.

## 1. Introduction

### 1.1 Spanish in contact with Valencian

This study aims to understand the attitudes toward morphosyntactic variation caused by the influence of Valencian in contact with Spanish in the province of Castellón. We use the matched-guise imitative technique to determine the perceptions with regards to morphosyntactic variation in a specific speech community (Lambert et al. 1960). This perception study complements the analysis of Blas Arroyo (1993), in which the author identified the sociolinguistic factors that condition morphosyntactic variation. Originally, matched-guise studies were used to comprehend the relationship between attitudes and two dialects or languages. Following Díaz-Campos & Killam (2012), we modified the matched-guise method

to observe the effect of variation in specific linguistic features in the speakers' attitudes. Few attitudinal studies in Spanish analyze the variability in perceptions (Díaz-Campos & Killam 2012; González Martínez & Blas Arroyo 2011; Wolck 1973; Woolard 1984; Woolard & Gahng 1990). The current study analyzes the relationship between morphosyntactic variation and participant attitudes in La Plana region (Castellón), which contributes to sociolinguistic studies by means of a matched-guise test to observe attitudes toward variation of the locative prepositions, partitive *de* and expletive *que* in interrogative sentences in Castellón. A strategy proposed by Labov (1972) is used to determine if this variation represents change in progress: the problem of evaluation in the understanding of subjective correlates in the observed changes.

While widely acknowledged as a problematic concept, Campbell-Kibler (2006) claims that the term 'prestige' refers to the attributes in the level of instruction, competence and correctness of language, while the term 'solidarity' refers to friendship, reliability or aesthetic qualities. Assigning prestige to one variant or another in the field of sociolinguistics is generally based on its social stratification. However, perception studies that verify how a variant is evaluated by its speakers are lacking (Thomas 2002). For instance, in the speech of the Valencian Community (Spain), cases of linguistic interference exist, due to its identity as a bilingual community in which Spanish is in contact with Valencian, a variety of the Catalan language. As a result of this contact situation, there is an influence of Valencian in these speakers' Spanish, affecting morphosyntactic and phonological systems. Blas Arroyo (1993) identified several examples of morphosyntactic interference in the speech of speakers from the Campanar district (Valencia). These include the locative prepositions *en/a* "in/at", as in *Estamos en/a casa* "we are in/at home"; the partitive use of the preposition *de* "of", as in *Me gusta esta camisa, pero las hay de/ø mejores* "I like this shirt, but there are ø better ones"; and the expletive *que* "that", as in *¿Que/ø tienes frío?* "ø Are you cold?" These variants are subject to social stratification, and they can be classified as 'standard' and 'non-standard' variants (Blas Arroyo 2011), or alternatively as variants that follow prescriptive grammar and variants that show interference from Valencian. In this study, regression coefficients were calculated for each independent variable using the results obtained in an acceptability test, an interference test and a linguistic insecurity test. In this manner, the authors found that socioeconomic class and level of instruction were the most significant independent variables conditioning the selection of variants of the dependent variable.

## 1.2 Valencian context

As Blas Arroyo (1993) highlights, certain features in Valencian grammar appear in the Spanish spoken by bilinguals in the Valencian Community. These features

include the locative prepositions, the partitive use of the preposition *de* and the expletive *que* in interrogative sentences.

### 1.2.1 Locative prepositions

The prepositions *a* and *en* “at” and “in” in Valencian correspond to the same prepositions in Spanish in terms of phonological shape, but there are differences in their usage. In Valencian, *a* and *en* uses can overlap. Both prepositions are used to express location and direction, but there is a tendency to use *a* for both location and direction when referring to a physical place. The preposition *en* usually appears before the words *un* “a” and *algún* “some” and the demonstrative pronouns *aquest* “this”, *aqueix* “that” and *aquell* “that (with a further distance)”. Wheeler et al. (2006) describe the basic pattern of use of these prepositions, distinguishing between *en* for location and *a* for direction (see Table 1). Nevertheless, the use of *a* to express location in all these contexts has an extended use in the province of Castellón.

Table 1. Prepositions *a* and *en* (Wheeler et al. 2006)

Before	Location <i>en</i>	Direction <i>a</i>
Noun	<i>No trobaren allotjament ni en hotels, ni en pensions, ni en cases particulars</i> “No encontraron alojamiento ni en hoteles, ni en pensiones, ni en casas particulares” “They didn’t find accommodation in hotels, hostels or private houses”	<i>Anem a sales diferents</i> “Vamos a salas diferentes” “We go to different rooms”
Adjective	<i>Viuen en grans edificis</i> “Viven en grandes edificios” “They live in big buildings”	<i>Viatgem a llunyanes terres</i> “Viajamos a tierras lejanas” “We travel to far lands”
Numeral	<i>He estudiat en tres universitats</i> “He estudiado en tres universidades” “I have studied at three universities”	<i>Els enviem a tres universitats</i> “Los enviamos a tres universidades” “We send them to three universities”
Quantifier/ indefinite	<i>Això passa en molts llocs</i> “Eso ocurre en muchos lugares” “That occurs in many places”	<i>Pujarem a quasi tots els pisos</i> “Subiremos a casi todos los pisos” “We will go up to most of the floors”
Relative pronoun	<i>El pis en el qual vivia</i> “El piso en el cual vivía” “The floor on which he/she lived”	<i>La terrassa a la qual hem pujat</i> “La terraza a la cual hemos subido” “The terrace to which we went up”

### 1.2.2 The partitive use of the preposition *de*

The preposition *de* is used with partitive value in the speech of some bilinguals in sentences such as *Me gusta esa camisa, pero las hay Ø mejores* “I like that shirt, but there are better ones”. This sentence adapts to the Valencian construction *n’hi ha de millors* (*me gusta esa camisa, pero las hay de mejores*). Wheeler et al. (2006) indicate that this represents a fixed idiomatic expression in Valencian. Speakers import into Spanish the tendency of including *de* before any adjectives that refer to the antecedent corresponding to the pronoun *en* (*en* is contracted to *n’* before a vowel or *h* + vowel).<sup>1</sup>

### 1.2.3 Expletive *que* in interrogative sentences

In spoken Valencian, the unstressed expletive *que* precedes short questions. It is optional, and its use does not indicate any change in emphasis or meaning, whereas in Spanish it does. Here we provide three examples:

- (a) *Hi ha ningú a la casa?* = *Que hi ha ningú a la casa?* (¿Ø/*Que* *hay alguien en casa?* “Is there anyone at home?”)
- (b) *Has pagat la factura del gas?* = *Que has pagat la factura del gas?* (¿Ø/*Qué* *has pagado la factura del gas?* “Did you pay the gas bill?”)
- (c) *No volem venir-hi?* = *Que no volem venir-hi?* (¿Ø/*Que* *no queréis venir?* “Don’t you want to come?”)

The current study will follow this organization: first, previous research will be linked to the morphosyntactic variables already mentioned; second, predictions and hypotheses will be posed; next, the methodology employed will be explained, followed by a presentation of the results; finally, we provide some conclusions, discuss limitations of the present study and suggest areas for further research.

---

1. In Valencian, the adverbial pronoun *en* exists, which represents the construction *de* + a Noun Phrase. This study focuses on two specific contexts. First, *en* can represent a Noun Phrase and have a partitive meaning. Since this type of pronoun does not exist in Spanish, it is normally represented with Ø in Spanish syntax (Bosque & Gutiérrez-Rexach 2008): *Ha vingut algun client?* –No, no *n’ha vingut cap encara* “Have any customers come? –No, no Ø clients have come yet”. Moreover, the preposition *de* precedes a noun when it is repeated before or after the verb: *De paciència no en té gens* “s/he has no Ø patience”.

## 2. Previous research

### 2.1 Language contact

The simplest definition for language contact (see Ramírez-Cruz, this volume) is the use of more than one language in a particular place (Thomason 2001). Recently, Matras (2009: i) has defined language contact as the situation “when speakers of different languages interact and their languages influence each other”. Since Weinreich’s (1953) studies on bilingualism, researchers have developed the concept of ‘interference’, but a consensus has not been reached concerning its definition (Blas Arroyo 1991; 1993). First, interference had negative connotations: errors that occur in the learning of a second language that reflect one’s acquisition of a first language and which are not found in the normal development of that second language as a native tongue (McLaughlin 1984). However, in recent years this concept has changed, and several linguists have proposed that interference is not a random phenomenon but something rather common that occurs frequently in bilingual communities (Blas Arroyo 1991, 1993, 1999; Gómez Molina 1984). Consequently, this phenomenon should not be understood in terms of deviations from norms but rather as a different grammar which gives place to concepts such as ‘grammaticality’ and ‘acceptability’ (Blas Arroyo 1993). Still, interference needs to be defined according to the specific situation and the languages that are in contact. At this point, it is necessary to mention that varieties such as Valencian and Catalan are used in all contexts, so these dialects cannot be defined as dialects of a minority language in the regions where they are spoken. Nevertheless, few specialists have focused on interference in Valencian-Spanish bilingual areas (Blas Arroyo 1991, 1993, 1999, 2011; Gimeno-Menéndez & Gómez Molina 2007; Gómez Molina 1984).

Following Blas Arroyo (2011), Valencian influence on Spanish can be measured. Certain forms can be clearly and directly attributed to Valencian. These variants are the basis of many stereotypes and are the most salient forms among speakers from outside this bilingual region. This type of interference includes calques from grammatical structures in Valencian, such as partitive *de* and the use of *a/en* with its locative function. However, other interference examples are more similar to Spanish and, therefore, are not obvious cases of interference. In these cases, bilingual speakers utilize linguistic forms that take advantage of the linguistic aspects shared by both grammars. The expletive *que* is a good example since Spanish has similar variants.

### 2.2 Attitudes

The concept of ‘attitude’ is known for the lack of agreement in its operationalization during the last few years. However, several linguists employ this term to

differentiate between favorable and unfavorable reactions toward certain phenomena (Blas Arroyo 1999; Sarnoff 1966). Attitudinal research in variationist sociolinguistics is scarce (Garrett et al. 2003; Giles & Powesland 1975; Rubino 2006; Von Gleich & Wölck 1994; Wölck 1973). Lafford (1986) is an example of an analysis of the prestige variable in the context of the Spanish language, which focuses on determining the values attributed to the different variants of /-s/ in syllable-final. In this case, the retention of the phoneme designates a more prestigious variety, the [h] (aspiration) is a neutral variant and the remaining variants are associated with lower prestige. Ros (1982) analyzes linguistic attitudes in the Valencian bilingual community. She acknowledges that in the 80s use of Spanish implied a higher socioeconomic class and educational level. Nevertheless, she finds a difference toward the variation of these values: the status of Valencian increases and is predicted to continue increasing in prestige in the coming years. In fact, Gómez Molina (1998) confirms that Valencian has recently been seen more positively, even surpassing Spanish in some contexts where the national language was considered the most prestigious variety (however, the opposite is observed in Casesnoves & Sankoff 2004). In Blas Arroyo (1995) and Gómez Molina (1998), it is demonstrated that attitudes vary according to linguistic and sociodemographic characteristics, observing that women from middle sociocultural levels evaluate Spanish guise speakers more positively while second generation participants from higher sociocultural classes attribute more positive values to Valencian guise speakers. One of the latest studies focuses on the region of Els Ports (Castellón) and analyzes the prestige of Valencian and Spanish (González Martínez & Blas Arroyo 2011). Their results suggest that Valencian is the less prestigious variety while Spanish is more prestigious. However, younger participants favor the usage of Valencian with respect to previous years. In the present study, we analyze the use of some morphosyntactic constructions in Spanish which are affected by Valencian influence, conducting a much more specific study, which will be explained below.

### 2.3 The matched-guise technique

The matched-guise technique emerged in Canada with the experiments by Lambert and his associates (1960) in the late 50s and early 60s. The first time this methodology was applied was in Lambert et al. (1960), which analyzed attitudes toward bilingualism. The technique consists of preparing various recordings in which a bilingual speaker reads the same passage in two different languages. Participants listen to these recordings and believe they are evaluating different guise speakers (Lambert 1967). These participants evaluate the personality of guise speakers according to a differential semantic scale in which they behave as judges of the recorded segments (Osgood et al. 1957). Thus, the matched-guise technique

serves to examine the languages or language varieties that coexist in a specific area. According to Blas Arroyo (2005), the objective of this technique is to observe and evaluate the psychosocial features associated with ‘guise speakers’ and relate this to the languages that are being employed. In the pioneer research about matched-guise, Lambert et al. (1960) use a sample of 64 English dominant participants and 66 French dominant participants who evaluate a bilingual speaker in both languages according to a 6-point differential semantic scale. The semantic differential scale “is a way to evaluate the emotional or subjective reactions from speakers concerning lexical elements with the goal to describe the affective dimensions in the organization of concepts in a linguistic variety” (Hernández Campoy & Almeida 2005). In Lambert et al.’s (1960) study, both participant pools favored the English guise in the success and solidarity scales. These findings demonstrated that the French minority group in Canada had attitudes similar to the stereotypes about Canadian French speakers held by the English majority group.

Díaz-Campos & Killam (2012) used matched-guise to analyze the attitudes toward variation in the production of one phoneme in a speech community. One of the main objectives in this study was to observe if the attitudes matched the conclusions of previous studies about stigmatized variants. In particular, they measured the attitudes toward elision and retention of intervocalic /d/ and /r/ in coda position in the speech of Caracas, Venezuela. They measured attitudes according to four attributes: intelligence, professional capacity, attractiveness and friendliness. For /r/ in coda position, two females read two versions of four sentences (one sentence retained the phoneme and the other elided it). Also, two males read two versions of four sentences for the retention/elision of intervocalic /d/. Díaz-Campos & Killam (2012) found that results did not correspond directly to the conclusions from previous studies regarding the production of intervocalic /d/ and /r/ in coda position. According to these studies, retention was the prestigious variant. However, researchers found that participants assigned a neutral evaluation to the retention and elision of intervocalic /d/ and a positive evaluation to /r/ retention. The authors confirmed that production studies indicate stratification of variants but that they do not show specific attitudes of speakers toward these variants.

### 3. Hypotheses and predictions

According to previous research, we expect to find support for Wolfram & Fasold’s (1974) claim: “Socially prestigious variants are the ones adopted by the high status group as linguistic indicators of their social status, while stigmatized variants are associated with lower class groups.” If their claim is supported, listeners in our study should assign more positive evaluations to ‘standard’ variants and less



positive values to ‘non-standard’ variants. As Blas Arroyo (2011) mentions, interference can be conceived as a scale. The variants partitive *de* and locative *a* represent the most notorious interference examples. In contrast, there are cases of interference that are less salient, such as expletive *que* in interrogative sentences. Therefore, we expect to find a major distinction between the positive and negative evaluations of partitive *de* and locative *a*, while we anticipate observing a more neutral distinction in the case of expletive *que*.

Previous studies of attitudes in the Valencian Community focus on how speakers evaluate Valencian in comparison to Spanish. To our knowledge, no studies have attempted to measure attitudes toward morphosyntactic variation caused by the contact of Spanish with Valencian. However, we can still formulate some predictions based on the previous literature. Although Gómez Molina (1998) demonstrates that the use of Valencian has been extended to new contexts in which Spanish was the prestige language, Casesnoves & Sankoff (2004) and González Martínez & Blas Arroyo (2011) find that Spanish is predominantly the more prestigious tongue. Therefore, we expect that listeners will assign a higher prestige to ‘standard’ variants. Furthermore, it is also possible that listeners will penalize guise speakers who show influence from Valencian, as this participant behavior occurred in Woolard (1984) and Woolard & Gahng (1990).

Regarding the solidarity sphere, previous studies demonstrate that participants who identify themselves as nationalists assign higher solidarity values to Valencian (Casesnoves & Sankoff 2004). Furthermore, younger participants tend to evaluate Valencian in a positive manner with respect to social attractiveness (González Martínez & Blas Arroyo 2011). If participants evaluate the variants that show some influence from Valencian in this same way, it can be hypothesized that participants will associate the Valencian identity with ‘non-standard’ variants. Therefore, it is anticipated that more positive evaluations for the ‘non-standard’ variants regarding solidarity and/or personal attractiveness.

## 4. Methodology

### 4.1 Participants

In this research, all participants were born and have lived most of their lives in the regions of La Plana, Castellón. La Plana is further subdivided in two smaller regions: La Plana Alta and La Plana Baixa, which are located in the southeast area of Castellón. Castellón de la Plana, Almazora, Benicasim, Oropesa del Mar, Torrelblanca, Borriol, Vall d’Alba, Cabanes, San Juan de Moró, Cuevas del Vinromá, Villafamés, Benlloch, Puebla Tornesa, Sierra Engarcerán, Villanueva de Alcolea, Torre Endoménech and Serratella are the cities and towns that belong to La Plana

Alta, which has a total of about 250,000 inhabitants. Villarreal, Burriana, Vall de Uxó, Onda, Nules, Moncófar, Almenara, Bechí, Alquerías del Niño Perdido, Vil-lavieja, Chilches, Artana, Ribesalbes, La Llosa, Eslida, Tales, Alfondiguilla, Sueras, Alcudia de Veo and Ahín are the municipalities found within La Plana Baja, approximately 200,000 people.

The four guise speakers who took part in the recordings of minimal pairs belong to the same socioeconomic level and generation and have similar professions. Roberto<sup>2</sup> is 36 years old and is the manager of a tile company. Esteban is 29 years old and is an administrator for a private company. Andrea is 29 years old and is in charge of a local Red Cross headquarters. Carla is 25 years old and cares for the elderly in a nursing home in Caux, France. All the guise speakers belong to the middle socioeconomic class and have achieved vocational school education. Clopper et al. (2005) found that the gender of guise speakers did not affect a correct categorization of dialects. To corroborate this finding in the present study, we include two males and two females, analyzing the gender effect in listeners' evaluations.

This study also includes 31 participants (13 males and 18 females) who evaluated the recorded sentences produced by the guise speakers mentioned above. Guise speakers and listeners do not know each other and have no contact, thereby avoiding the possibility that listeners are familiar with the guise speakers' voices. The listeners belong to different generations, have varying levels of education and have been bilingual in Spanish and Valencian since infancy. The researchers made use of a background questionnaire to ensure that at least one parent spoke Valencian in the household and, therefore, bilingualism was self-reported by guise speakers by means of indicating the language that their parents used with them since they were born. Most participants were contacted by one of the authors via e-mail, while others completed the study through a post on a social network.

## 4.2 Instrument

Three instruments were used to determine the attitudinal evaluations assigned by listeners: one for partitive *de*, another for locative *a/en* and finally one for expletive *que*. For partitive *de*, each speaker recorded two sentences with *de* and two without *de*:

- (1) *Me gusta esta camisa, pero las hay de mejores* ("I like that shirt, but there are better ones").

---

2. Pseudonyms are used for all participants.

- (2) *Se coge dos o las que quiera de manzanas* (“you can take two apples or as many as you want”).
- (3) *Me gusta esta camisa, pero las hay mejores* (“I like that shirt, but there are Ø better ones”).
- (4) *Se coge dos o las manzanas que quiera* (“you can take two apples or as many Ø as you want”).

Sentences with *de* represent the Spanish variety influenced by contact with Valencian, and sentences without *de* show the canonical Spanish variety. For locative *a/en* perception, each speaker recorded a sentence with *a* and another with *en*:

- (1) *Los niños están a la playa* (“the children are at the beach”).
- (2) *Pasamos todo el día al mercado* (“we spent all day at the market”).
- (3) *Los niños están en la playa* (“the children are in the beach”).
- (4) *Pasamos todo el día en el mercado* (“we spent all day in the market”).

Sentences with *a* represent the Spanish variety influenced by contact with Valencian, and sentences with *en* are examples of the Standard Spanish variety. For the perception of expletive *que*, each speaker recorded the following sentences:

- (1) *¿Que tienes frío?* (“Are you cold?”)
- (2) *¿Que no quieres ir al cine conmigo?* (“Don’t you want to come to the movies with me?”)
- (3) *¿Tienes frío?* (“Ø Are you cold?”)
- (4) *¿No quieres ir al cine conmigo?* (“Ø Don’t you want to come to the movies with me?”)

Sentences with *que* show the Spanish variety influenced by Valencian, and sentences without *que* reflect the canonical Spanish variety.

In analyzing the perception of morphosyntactic variation in these three contexts, voice quality and gender effects of the guise speakers must be taken into account. Campbell-Kibler (2006) observed that lexical items in recordings might influence the listeners’ evaluations. To control for this effect, listeners evaluate these different constructions in otherwise identical sentences. In all the sentences, only the lexical item associated with each morphosyntactic variable changes. Indeed, each sentence has two pairs of adjectives (“bad/good”, “condescending/humble”, “poor/rich” and “non-intelligent/intelligent” being the adjective binomials used). This methodology allows us to determine if both variants can overcome quality voice effects among these four guise speakers and it offers the possibility to compare differences between them. Forty-eight sequences are included in this

analysis; details about these sequences appear in the Appendix. Thirty-one listeners evaluated 48 sentences and two pairs of adjectives per recording (see examples 1 and 2 below), obtaining a total of 2,976 evaluations and 992 for each variable:

- (1) Audio: “Los niños están a la playa”. Esta persona es:

1 2 3 4 5 6

No inteligente ... inteligente

1 2 3 4 5 6

Mala ... buena

- (2) Audio: ¿Que tienes frío? Esta persona es:

1 2 3 4 5 6

Condescendiente ... humilde

1 2 3 4 5 6

Pobre ... rica

The order of sequences ensures that listeners compare each guise speaker with a speaker of the same gender and two guise speakers of the opposite gender (one with the canonical use and another with the non-canonical use). The sentences were randomized for presentation (see Appendix).

Apart from these recordings, the instrument also includes a matched-guise questionnaire with a differential scale from 1 to 6. This scale allows listeners to evaluate the psychosocial features that listeners assign to the different variants. Carranza (1982) categorizes attitudes in two dimensions: solidarity and prestige. In the same way, González Martínez & Blas Arroyo (2011) describe three spheres for these attitudes: personal attractiveness, social attractiveness and socioeconomic status. The current study groups adjectives for the matched-guise according to personal attractiveness and socioeconomic status. The first refers to solidarity and the latter to prestige, as described in Carranza (1982). For the personal attractiveness sphere, the binomials “bad/good” and “condescending/humble” are used. For the socioeconomic status sphere, the binomials “poor/rich” and “non-intelligent/intelligent” are employed. The matched-guise questionnaire was completed through *Quia*, an educational website that allows users to create questionnaires among other materials. In this portal, listeners were asked to evaluate each recording from 1 to 6 (1 represents the most negative value and 6 the most positive value).

## 5. Results

To analyze the evaluations’ means, the software SPSS (17.0) was used, comparing the mean of the Attitudes dependent variable while controlling the fixed and random effects for each of the different morphosyntactic variables in a mixed-model effects analysis (Díaz-Campos & Killam 2011). Consequently, the speaker is coded

as a random effect while gender is coded as a fixed effect. We present results in this order: First, we describe participants' general evaluations about the recorded stimuli in order to identify differences among variants with Valencian interference and variants adhering to more standard Spanish norms. After that, evaluations according to the different guise speakers are shown to determine if voice quality and/or gender effects exist. Finally, evaluations following the personal attractiveness category (which includes the binomials “bad/good” and “condescending/humble”) and socioeconomic status category (which consists of the binomials “poor/rich” and “non-intelligent/intelligent”) spheres are utilized.

In Figures 1, 2 and 3, general evaluations about the variants locative *a/en*, partitive *de* and expletive *que* are presented. Thirty-one listeners evaluated a total of 12 sentences per speaker according to two different scales with different adjectives: four sentences including the distinction between the prepositions *a/en*, four sentences with/without the use of partitive *de* and four sentences with/without the use of expletive *que*. In total, 96 evaluations were collected for each of the guise speakers: Carla, Roberto, Andrea and Esteban.

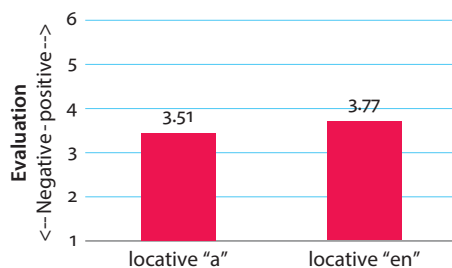


Figure 1. Evaluations for locative *a* and *en*

Evaluations for locative *a* have an average score of 3.51 while evaluations of sentences containing locative *en* received an average score of 3.77 (see Figure 1). It is important to remember that, in the scale provided to participants, 1 is the most negative value and 6 is the most positive. The distinction between these locative prepositions determines that locative *en* obtains more positive evaluations according to the adjectives employed for the personal attractiveness and socioeconomic status spheres. This tendency is highly significant in the mixed model analysis  $F(1, 216) = 11.398, p = 0.001$ . A 3.70 average was obtained for the use of partitive *de* (see Figure 2), while a lower average was obtained for the sentences without partitive *de*: 3.65. Therefore, the presence of partitive *de* entails more positive evaluations among listeners. However, the difference between these averages is not significant in the statistical analysis  $F(1, 213) = 0.515, p = 0.474$ . The statistical analysis shows that the difference between the variants for expletive *que* is not significant either  $F$

(1, 213) = 0.612,  $p = 0.435$  (see Figure 3). The average score obtained for those sentences containing expletive *que* was 3.72, showing a minimal difference from the average score of 3.77 assigned to the sentences without expletive *que*.

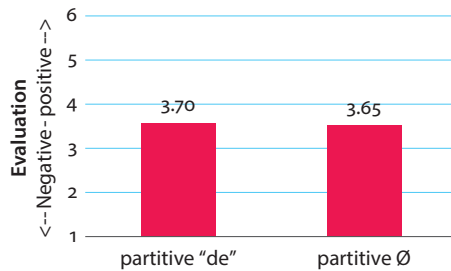


Figure 2. Evaluations for partitive *de*

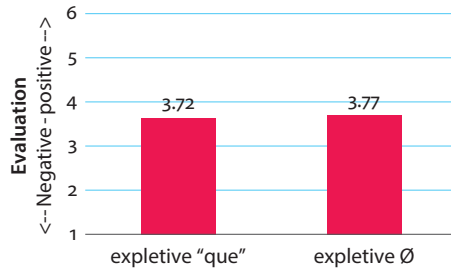


Figure 3. Evaluations for expletive *que*

With respect to the speaker's voice quality, Carla received more positive evaluations than the rest of the guise speakers, although this difference is not statistically significant. Carla's average score is 3.75, while Andrea's is 3.59, Roberto's is 3.67 and Esteban receives an average score of 3.74. If all the averages are compared, there are no significant differences between any of the guise speakers (Andrea vs. Esteban  $p = 0.088$ ; Andrea vs. Roberto  $p = 1.000$ ; Esteban vs. Roberto  $p = 1.000$ ; Carla vs. Andrea  $p = 0.058$ ; Carla vs. Esteban  $p = 1.000$ ; Carla vs. Roberto  $p = 1.000$ ). Thus, it is clear that no voice quality effect existed in this study, and this was controlled by means of providing clear instructions to guise speakers about the recording process. In particular, they were asked to record the sentences at a normal pace with a neutral tone and with minimal background noise. The gender of the guise speakers is also not a significant predictor for participants' evaluations (see Figure 4). The average score for female guise speakers is 3.67, while the males receive an only slightly higher average score of 3.70. Consequently, guise speaker gender does not affect listeners' evaluations of these sentences  $F(1, 2) = 0.153$ ,  $p = 0.734$ .

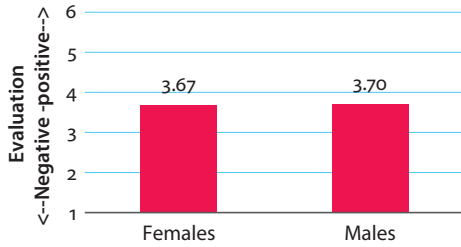


Figure 4. Evaluations according to guise speakers' gender

With respect to the evaluations obtained for each variant according to the speaker (see Figure 5), only some variants show statistically significant differences. In this case, for the distinction between locative *a* and *en*, the sentences produced by Carla and Roberto have a significant difference in their evaluations (Carla  $F(1, 30) = 5.595, p = 0.025$ ; Roberto  $F(1, 30) = 5.638, p = 0.024$ ) and, in general, the use of locative *en* is evaluated with higher values than locative *a*. On the other hand, the differences between evaluations for Andrea and Esteban are not significant (Andrea  $F(1, 30) = 0.907, p = 0.349$ ; Esteban  $F(1, 30) = 2.352, p = 0.136$ ). Furthermore, the overall difference between *a* and *en* is not significant  $F(1, 210) = 0.321, p = 0.810$ ).

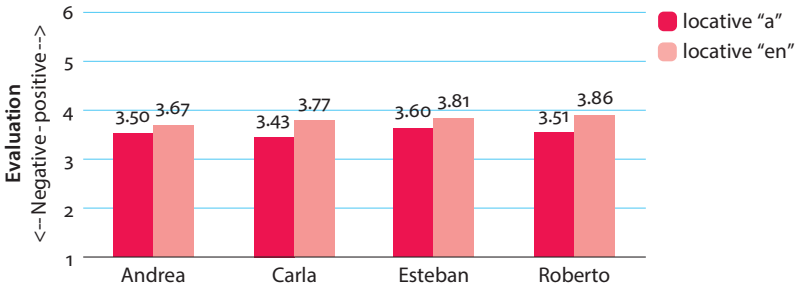


Figure 5. Evaluations for *a* and *en* locative variants per speaker

With respect to the presence and absence of partitive *de* (see Figure 6), none of the guise speakers differ significantly from the others (Andrea  $F(1, 30) = 0.229, p = 0.636$ ; Carla  $F(1, 30) = 3.552, p = 0.069$ ; Esteban  $F(1, 30) = 0.294, p = 0.592$ ; Roberto  $F(1, 30) = 1.212, p = 0.280$ ). In other words, evaluations among these two variants are similar across guise speakers. Although in all cases this trend (which is not significant) is followed, only in the productions by Roberto is the absence of partitive *de* evaluated more positively than its presence. Overall, the difference between the presence and absence of *de* is not significant  $F(1, 210) = 1.502, p = 0.215$ .

Concerning the presence and absence of expletive *que* (see Figure 7), none of the differences in evaluation by speaker are significant (Andrea  $F(1, 30) = 0.914$ ,  $p = 0.347$ ; Carla  $F(1, 30) = 0.305$ ,  $p = 0.585$ ; Esteban  $F(1, 30) = 0.646$ ,  $p = 0.428$ ; Roberto  $F(1, 30) = 0.045$ ,  $p = 0.833$ ). However, the absence of expletive *que* is evaluated more positively than its presence, except in Carla's case. In the same manner, the overall difference between the presence and absence of expletive *que* is not significant ( $F(1, 210) = 0.565$ ,  $p = 0.639$ ).

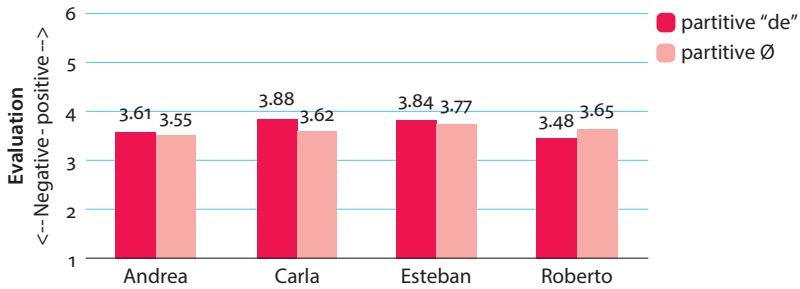


Figure 6. Evaluations for partitive *de* variants per guise speaker

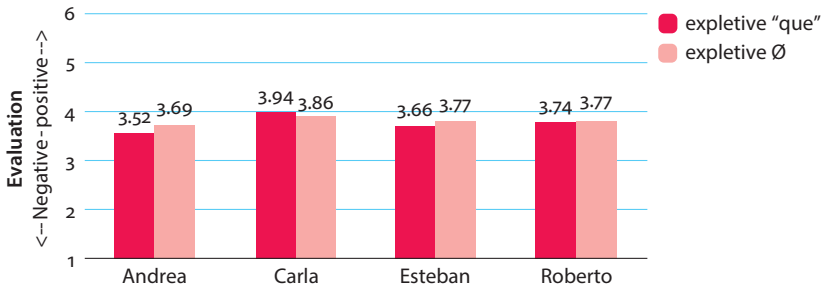


Figure 7. Evaluations for *que* expletive variants per speaker

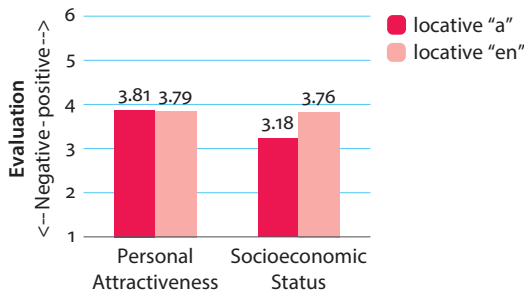


Figure 8. Evaluations for locative *a* and *en* variants by adjective



Last, the results with respect to the Socioeconomic Status and Personal Attractiveness spheres are shown. In the case of the prepositions *a* and *en* (see Figure 8), listeners provided more positive evaluations for the use of locative *en* in the Socioeconomic Status. In this respect, Socioeconomic Status is the only statistically significant sphere ( $F(1, 213) = 28.446, p = 0.000$ ), while Personal Attractiveness is not significant ( $F(1, 216) = 2.189, p = 0.140$ ). This variable is the only one showing a significant difference between the use of the two variants in the Personal Attractiveness and Socioeconomic Status spheres ( $F(1, 462) = 13.907, p = 0.000$ ).

Regarding partitive *de*, listeners have evaluated both spheres in a similar way for the presence and absence of this preposition (see Figure 9). Within each sphere, there is not a significant difference between the two variants in the analysis, with  $F(1, 213) = 0.224, p = 0.637$  for the Socioeconomic Status sphere and  $F(1, 213) = 0.359, p = 0.550$  for the Personal Attractiveness sphere. On the other hand, listeners have evaluated fairly positively a variant of expletive *que* according to a specific sphere, although their values are not statistically significant with respect to Personal Attractiveness ( $F(1, 213) = 0.282, p = 0.596$ ) or Socioeconomic Status ( $F(1, 213) = 3.165, p = 0.077$ ) (see Figure 10). Consequently, it can be determined that participants in this study, who live in the province of Castellón, assign more attractive personal values and a higher socioeconomic level to guise speakers when they use some variants commonly found in the bilingual contact situation between Spanish and Valencian.

In conclusion, no significant differences were found between expletive *que* and the 'standard'  $\emptyset$  expletive and between partitive *de* and its absence, but a significant difference was observed between locative *a* and *en*. As will be discussed below, these findings may be related to the level of integration of these 'non-standard' variants in the Spanish spoken in this area; it may be that the variant locative *a* does not yet fit completely within this linguistic system. Finally, no gender or voice effects were detected, reinforcing the validity of the mixed model analysis to discern speakers' attitudes in a bilingual community.

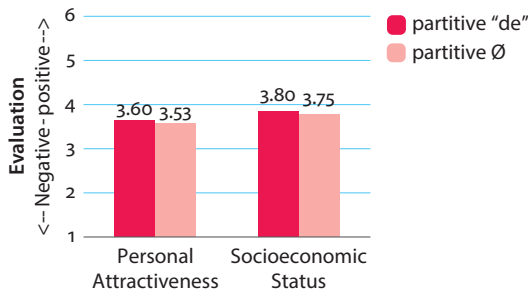


Figure 9. Evaluations for partitive *de* variants by adjective

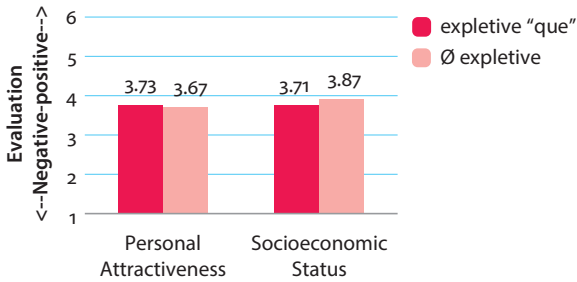


Figure 10. Evaluations for expletive *que* variants by adjective

## 6. Discussion

Our results show the importance of using perception instruments to elucidate prestige values assigned to variants that are stratified sociolinguistically. Although no statistical differences were found between expletive *que* and Ø expletive or between partitive *de* and Ø partitive, statistical differences were obtained between locative *a* and *en*. Results related to expletive *que* confirm our hypothesis; according to the interference scale suggested by Blas Arroyo (2011), expletive *que* can be seen as a case of non-obvious interference, given that several rather similar constructions exist in Spanish. Furthermore, it seems that listeners do not perceive a difference between the variants partitive *de* and Ø. Nevertheless, the locative *a* variant is understood as a completely different case, since it represents a clear case of interference from Valencian.

The significant difference between locative *en* and *a* suggests that the 'standard' variant *en* is the prestigious variant. Moreover, expletive *que* follows this same pattern. That is, the canonical variant Ø expletive receives more positive evaluations than the expletive *que* variant, although this difference is not significant. These results imply that there are differences in their sociolinguistic meaning. In this case, the locative *a* variant is more prominent in terms of perception, which is in line with Blas Arroyo (1993). From the variants he analyzed, locative *a* was shown to belong to the group of variants causing the greatest heterogeneity among the population sampled. Indeed, Blas Arroyo (1993) noticed that locative *a* stood out for its low frequency and notable dispersion in society. He concluded that locative *a* is one of the most stigmatized variants in this province. This connection with our results implies that locative *a* is a vernacular variant in spontaneous speech. However, the partitive *de* variant does not show the same pattern as locative *a* because partitive *de* obtained more positive overall evaluations than partitive Ø. This could indicate a positive response to Valencian identity, but since this difference is not significant it is more likely that this variant is not perceived as

‘non-standard’ in this region. Therefore, the partitive *de* variant does not seem to be stigmatized among this population.

With the exception of the variant partitive *de*, global variants’ evaluations show that ‘non-standard’ variants received more negative values and, most importantly, they reflect the social stratification suggested in the production study in Blas Arroyo (1993). This dynamic becomes clearer when analyzing the evaluations according to the two attitudinal spheres of Personal Attractiveness and Socioeconomic Status. Different results were obtained for each variant. In general, guise speakers who produce the ‘standard’ locative *en* variants and expletive  $\emptyset$  are perceived as intelligent and rich. Nevertheless, the Personal Attractiveness sphere showed the opposite trend; the ‘non-standard’ variants received more positive evaluations, though this was not statistically significant. Díaz-Campos & Killam (2012) affirmed that it is expected that attitudes with regard to intelligence and professionalism be more salient when evaluating linguistic behavior. This notion is related to the claim made in Clopper et al. (2005): speakers associate a ‘standard’ variant with intelligence while ‘non-standard’ variants receive more positive evaluations in terms of solidarity. Consequently, listeners that speak a ‘non-standard’ variety often assign more positive evaluations to a speaker who speaks the same variety. In this regard, our results suggest that listeners distribute features for the Personal Attractiveness sphere with the variety of the language that is spoken in their region, which supports Lambert et al.’s (1960) findings in their matched-guise study of English and French in Canada. In spite of the different research purposes of Lambert et al. (1960), the similarities between this and the present study are clear, given that our research isolated examples of linguistic variation within the same variety of Spanish.

Although evaluations for expletive *que* followed the same tendencies as locative *a*, the difference between variants’ averages is only significant for locative *a*. These results suggest a general acceptance of the expletive *que* variant and a certain degree of stigmatization of the locative *a* variant in this community. As previously mentioned, since *que* is grammatical in certain constructions in Spanish, it is possible that listeners do not realize that expletive *que* is an example of Valencian interference. However, partitive *de* does not follow the pattern observed with expletive  $\emptyset$  and locative *en*. The ‘standard’ variant partitive  $\emptyset$  obtained more negative evaluations for both spheres, though these differences were not significant. It is likely that partitive *de* is not consciously associated with Valencian interference, and it may entail some local prestige. This last interpretation is supported by our results, since partitive *de* obtained more positive overall evaluations.

An innovative aspect of the present study is the inclusion of female and male guise speakers in the matched-guise. This enabled us to determine if gender played a role in the perceptions of listeners. When combining the averages of females and

males, there was no significant difference, which indicates that gender does not have a significant impact on the perception of social features. Consequently, the methodology that has been utilized in this study seems appropriate in order to isolate the effect of morphosyntactic variants in listeners' evaluations. Results in this study help to understand how listeners perceive morphosyntactic variation as a result of intimate Valencian and Spanish contact. While Blas Arroyo (1993) found social stratification of the same morphosyntactic variables included in this study, the analysis performed here demonstrates the importance of complementing this research with perception studies. This way, we can better understand linguistic interference in bilingual communities, as well as how attitudes differ according to the morphosyntactic variation that arises in these contact situations.

## 7. Conclusions, limitations and future research

Our findings partially support the idea that socially prestigious variants are the ones adopted by speakers of the higher social classes (Wolfram & Fasold 1974). However, it is necessary to take into account the specific dynamics of the bilingual context in this study (see Riccelli, this volume). Our results suggest that expletive *que* and partitive *de* are not socially prominent, while locative *a* is. This phenomenon partially matches the interference scale in Blas Arroyo (2011). Furthermore, the results for the three 'non-standard' variants may imply that listeners express solidarity toward speakers from their own community. The fact that partitive *de* obtained more positive evaluations for the Socioeconomic Status sphere suggests that it might have a higher local prestige when compared to the locative *a* and expletive *que* variants, which received more negative evaluations in this sphere. Moreover, it seems that listeners are aware about the Valencian interference in the case of the locative *a* variant, given that it received significant negative evaluations in the Socioeconomic Status sphere. This finding makes clear how the linguistic system in this particular community is different from other regions in the peninsula due to its contact with Valencian. Studies such as this one are necessary in order to be able to determine how these varieties differ and what patterns are accepted by the members of the community as forms become integrated into the linguistic system. In our case, the preposition *a* does not seem to be as integrated in the speech of these speakers as the expletive *que* and partitive *de* variants.

A number of limitations that hindered the present study may help future researchers in their work on this topic. For instance, in this study, a sequence of 48 recordings was employed in order to receive evaluations for three morphosyntactic variables. We limited the number of adjectives per recording to reduce participants' workload. However, the questionnaire was still long, which may have

caused fatigue among participants and, consequently, could have affected the evaluations obtained. For future studies, employing a shorter perception questionnaire is recommended. Another limitation is the format of the questionnaire in *Quia*. Despite researchers' giving speakers the choice to complete the questionnaire in Valencian or in Spanish, some instructions in the questionnaire appeared in English by default, which may have caused some confusion. Moreover, guise speakers recorded the sentences in a natural context without the presence of the researchers, which caused some difficulties when ensuring similar prosody, intonation and stress. Ideally, these differences could be minimized by having one of the researchers record the guise speakers and ensure that recordings have a similar quality.

The present study is relevant to the analysis of the attitudes of a specific community with respect to spoken language. In this case, we find that a linguistic variety is different from other peninsular varieties due to the bilingual character of this specific region. Particularly, speakers from La Plana, Castellón, perceive the preposition *a* as a noticeable result of transfer from Valencian, while the variants expletive *que* and partitive *de* seem to be totally integrated into the language of this bilingual community. Also, the validity of the mixed model effects analysis is supported since no voice or gender effect was detected. This study helps identify ongoing processes in this variety of Spanish and determine which variants are socially prestigious and/or have been fully integrated into the linguistic system.

## References

- Blas Arroyo, José L. 1991. "Problemas teóricos en el estudio de la interferencia lingüística". *Revista Española de Lingüística* 2.265–289.
- Blas Arroyo, José L. 1993. *La interferencia lingüística en Valencia (dirección, catalán-español)*. Castellón: Universitat Jaume I.
- Blas Arroyo, José L. 1995. "La dirección de la interferencia lingüística a debate: A propósito de un caso de simplificación gramatical en el español por influencia del catalán". *Revue Roumaine de Linguistique* 40:5–6.30–56.
- Blas Arroyo, José L. 1999. "Las actitudes hacia la variación intradialectal en la sociolingüística hispánica". *Estudios filológicos* 34.47–72.
- Blas Arroyo, José L. 2005. *Sociolingüística del español: Desarrollos y perspectivas en el estudio de la lengua española en el contexto social*. Madrid: Cátedra.
- Blas Arroyo, José L. 2011. "Spanish in Contact with Catalan". *The Handbook of Hispanic Sociolinguistics* ed. by Manuel Díaz-Campos, 374–394. Malden, Mass.: Blackwell.  
doi:10.1002/9781444393446.ch18
- Bosque Muñoz, Ignacio & Javier Gutiérrez-Rexach. 2009. *Fundamentos de sintaxis formal*. Madrid: Ediciones AKAL.
- Campbell-Kibler, Katherine. 2006. *Listener Perceptions of Sociolinguistic Variables: The Case of (ING)*. Ph.D. dissertation, Stanford University.

- Carranza, Miguel A. 1982. "Attitudinal Research on Hispanic Language Varieties". *Attitudes toward Language Variation* ed. by Ellen Bouchard Ryan & Howard H. Giles, 63–84. London: Edward Arnold.
- Casesnoves Ferrer, Raquel & David Sankoff. 2004. "The Valencian Revival: Why Usage Lags behind Competence". *Language in Society* 33.1–31.
- Cooper, Robert L. & Joshua A. Fishman. 1975. "The Study of Language Attitudes". *Linguistics* 136.5–16.
- Clopper, Cynthia G., Brianna Conrey & David B. Pisoni. 2005. "Effects of Talker Gender on Dialect Categorization". *Journal of Language and Social Psychology* 24.182–206. doi:10.1177/0261927X05275741
- Díaz-Campos, Manuel & Jason Killam. 2012. "Assessing Language Attitudes through a Matched-Guise Experiment: The Case of Consonant Deletion in Venezuelan Spanish". *Hispania* 95.83–102.
- Fasold, Ralph 1984. *The Sociolinguistics of Society*. Oxford: Basil Blackwell.
- Garrett, Peter, Nikolas Coupland & Angie Williams. 2003. "Investigating Language Attitudes: Social Meanings of Dialect, Ethnicity and Performance". *International Journal of Applied Linguistics* 15:3.411–414.
- Giles, Howard & Peter F. Powesland. 1975. *Speech Style and Social Evaluation*. New York: Academic Press.
- Gimeno-Menéndez, Francisco & José R. Gómez Molina. 2007. "Spanish and Catalan in the Community of Valencia". *International Journal of the Sociology of Language* 184.95–107.
- González Martínez, Juan & José L. Blas Arroyo. 2011. "Estabilidad y dinamismo en las actitudes lingüísticas de una comunidad bilingüe española (Els Ports, Castellón)". *Hispania* 94.663–679.
- Gómez Molina, José R. 1984. "Actituds i usos lingüístics en la comunitat de parla de Sagunt (València)". *Cuadernos de Filología II: Studia Lingüística Hispánica* 3.75–104.
- Gómez Molina, José R. 1998. "Actituds lingüístiques a València: Avaluació de quatre varietats dialectals". *El contacte i la variació lingüística: Descripció i metodologia* ed. by Emili Boix, Nuria Alturo, Maria Rosa Lloret, Mercè Lorente & Lluís Payrató. Barcelona: Universitat de Barcelona.
- Hernández Campoy, Juan M. & Manuel Almeida. 2005. *Metodología de la investigación sociolingüística*. Albolote: Comares.
- Labov, William. 1972. "On the Mechanism of Linguistic Change". *Sociolinguistic Patterns* ed. by William Labov, 160–182. Philadelphia: University of Pennsylvania Press.
- Lafford, Barbara. 1986. "Valor diagnóstico-social del uso de ciertas variantes de /s/ en el español de Cartagena, Colombia". *Estudios sobre la fonología del español del Caribe* ed. by Rafael A. Núñez Cedeño, Iraset Páez & Jorge M. Guitart, 53–55. Caracas: La Casa de Bello.
- Lambert, Wallace 1976. "The Social Psychology of Bilingualism". *Journal of Social Issues* 23. 91–109. doi:10.1111/j.1540-4560.1967.tb00578.x
- Lambert, Wallace, Richard C. Hodgson, Robert C. Gardner & Samuel Fillenbaum. 1960. "Evaluational Reactions to Spoken Languages". *Journal of Abnormal and Social Psychology* 60:1.44–51. doi:10.1037/h0044430
- Mackey, William F. 1970. *Bilinguisme et contact des langues*. Paris: Klincksieck.
- Matras, Yaron 2009. *Language Contact*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511809873
- Matus-Mendoza, María L. 2004. "Assibilation of /-r/ and Migration Among Mexicans". *Language Variation and Change* 16.17–30. doi:10.1017/S0954394504161024

McLaughlin, Barry. 1984. *Second Language Acquisition in Childhood: Preschool Children*. Hillsdale, N.J.: Lawrence Erlbaum Associates.

Milroy, James & Lesley Milroy. 1985. *Authority*. London: Routledge.

Osgood, Charles, George Suci & Percy Tannenbaum. 1957. *The Measurement of Meaning*. Urbana: University of Illinois Press.

Ros, María. 1982. “Percepción y evaluación de los hablantes de cinco variedades lingüísticas”. *Estructura social al País Valencià* ed. by Rafael L. Ninyoles, 679–698. Valencia: Diputació Provincial.

Rubino, Antonio. 2006. “Linguistic Practices and Language Attitudes of Second-Generation Italo-Australians”. *International Journal of the Sociology of Language* 180.71–88.

Sarnoff, Irving. 1966. “Social Attitudes and the Resolution of Motivational Conflict”. *Attitudes* ed. by Marie Jahoda & Neil Warren, 279–284. Harmondsworth: Penguin.

SPSS for Windows. Rel. 17.0.2. 2010. Chicago: SPSS. Software.

Thomas, Erik R. 2002. “Sociophonetic Applications of Speech Perception Experiments”. *American Speech* 77.115–147. doi:10.1215/00031283-77-2-115

Thomason, Sarah G. 2001. *Language Contact*. Washington: Edinburgh University Press. doi:10.1016/Bo-08-043076-7/03032-1

Von Gleich, Uta & Wolfgang Wölck. 1994. “Changes in Language Use and Attitudes of Quechua-Spanish Bilinguals in Peru”. *Language in the Andes* ed. by Peter Cole, Gabriella Hermon & Mario Daniel Martín, 27–55. Newark, Del.: University of Delaware.

Weinreich, Uriel. 1953. *Languages in Contact. Findings and Problems*. New York: Publications of the Linguistic Circle of New York.

Wheeler, Max W., Alan Yates & Nicolau Dols. 2006. *Catalan: A Comprehensive Grammar*. New York: Routledge. doi:10.1016/Bo-08-044854-2/02246-X

Wolck, Wolfgang. 1973. “Attitudes toward Spanish and Quechua in Bilingual Peru”. *Language Attitudes: Current Trends and Prospects* ed. by Roger Shuy & Ralph W. Fasold, 129–147. Washington: Georgetown University Press.

Wolfram, Walter A. & Ralph W. Fasold. 1984. *The Study of Social Dialects in American English*. Ann Arbor, Mich.: Prentice Hall.

Woolard, Kathryn A. & Tae-Joong Gahng. 1990. “Changing Language Policies and Attitudes in Autonomous Catalonia”. *Language in Society* 19.311–330. doi:10.1017/S0047404500014536

Woolard, Kathryn A. 1984. “A Formal Measure of Language Attitudes in Barcelona”. *International Journal of the Sociology of Language* 47.63–71.

## Appendix

### Sequence of recordings

Sequence	Speaker	Sentence	Variant	Attitude
1	Roberto	<i>Los niños están a la playa</i>	locative <i>a</i>	Personal Attractiveness
2	Andrea	<i>Me gusta esta camisa, pero las hay de mejores</i>	partitive <i>de</i>	Personal Attractiveness
3	Esteban	<i>¿Que no quieres ir al cine conmigo?</i>	expletive <i>que</i>	Personal Attractiveness



## Appendix (Continued)

Sequence	Speaker	Sentence	Variant	Attitude
4	Carla	<i>Se coge dos o las que quiera de manzanas</i>	partitive <i>de</i>	Socioeconomic Status
5	Esteban	<i>Pasamos todo el día al mercado</i>	locative <i>a</i>	Socioeconomic Status
6	Andrea	<i>¿Que tienes frío?</i>	expletive <i>que</i>	Socioeconomic Status
7	Esteban	<i>Se coge dos o las que quiera de manzanas</i>	partitive <i>de</i>	Socioeconomic Status
8	Roberto	<i>¿Que tienes frío?</i>	expletive <i>que</i>	Socioeconomic Status
9	Carla	<i>Pasamos todo el día en el mercado</i>	locative <i>en</i>	Socioeconomic Status
10	Andrea	<i>Me gusta esta camisa, pero las hay mejores</i>	Ø partitive	Personal Attractiveness
11	Esteban	<i>¿No quieres ir al cine conmigo?</i>	Ø expletive	Personal Attractiveness
12	Roberto	<i>Me gusta esta camisa, pero las hay de mejores</i>	partitive <i>de</i>	Personal Attractiveness
13	Andrea	<i>Los niños están a la playa</i>	locative <i>a</i>	Personal Attractiveness
14	Esteban	<i>Pasamos todo el día en el mercado</i>	locative <i>en</i>	Socioeconomic Status
15	Roberto	<i>Los niños están en la playa</i>	locative <i>en</i>	Personal Attractiveness
16	Carla	<i>¿Que no quieres ir al cine conmigo?</i>	expletive <i>que</i>	Personal Attractiveness
17	Roberto	<i>¿Tienes frío?</i>	Ø expletive	Socioeconomic Status
18	Esteban	<i>Se coge dos o las manzanas que quiera</i>	Ø partitive	Socioeconomic Status
19	Andrea	<i>¿Tienes frío?</i>	Ø expletive	Socioeconomic Status
20	Carla	<i>Se coge dos o las manzanas que quiera</i>	Ø partitive	Socioeconomic Status
21	Andrea	<i>Los niños están en la playa</i>	locative <i>en</i>	Personal Attractiveness
22	Carla	<i>Pasamos todo el día al mercado</i>	locative <i>a</i>	Socioeconomic Status
23	Roberto	<i>Me gusta esta camisa, per las hay mejores</i>	Ø partitive	Personal Attractiveness
24	Carla	<i>¿No quieres ir al cine conmigo?</i>	Ø expletive	Personal Attractiveness
25	Esteban	<i>Los niños están a la playa</i>	locative <i>a</i>	Personal Attractiveness
26	Carla	<i>Me gusta esta camisa, pero las hay de mejores</i>	partitive <i>de</i>	Personal Attractiveness

(Continued)



## Appendix (Continued)

Sequence	Speaker	Sentence	Variant	Attitude
27	Roberto	<i>¿Que no quieres ir al cine conmigo?</i>	expletive <i>que</i>	Personal Attractiveness
28	Andrea	<i>Se coge dos o las que quiera de manzanas</i>	partitive <i>de</i>	Socioeconomic Status
29	Roberto	<i>Pasamos todo el día al mercado</i>	locative <i>a</i>	Socioeconomic Status
30	Carla	<i>¿Que tienes frío?</i>	expletive <i>que</i>	Socioeconomic Status
31	Roberto	<i>Se coge dos o las que quiera de manzanas</i>	partitive <i>de</i>	Socioeconomic Status
32	Esteban	<i>¿Que tienes frío?</i>	expletive <i>que</i>	Socioeconomic Status
33	Andrea	<i>Pasamos todo el día en el mercado</i>	locative <i>en</i>	Socioeconomic Status
34	Carla	<i>Me gusta esta camisa, per las hay mejores</i>	Ø partitive	Personal Attractiveness
35	Roberto	<i>¿No quieres ir al cine conmigo?</i>	Ø expletive	Personal Attractiveness
36	Esteban	<i>Me gusta esta camisa, pero las hay de mejores</i>	partitive <i>de</i>	Personal Attractiveness
37	Carla	<i>Los niños están a la playa</i>	locative <i>a</i>	Personal Attractiveness
38	Roberto	<i>Pasamos todo el día en el mercado</i>	locative <i>en</i>	Socioeconomic Status
39	Esteban	<i>Los niños están en la playa</i>	locative <i>en</i>	Personal Attractiveness
40	Andrea	<i>¿Que no quieres ir al cine conmigo?</i>	expletive <i>que</i>	Personal Attractiveness
41	Esteban	<i>¿Tienes frío?</i>	Ø expletive	Socioeconomic Status
42	Roberto	<i>Se coge dos o las manzanas que quiera</i>	Ø partitive	Socioeconomic Status
43	Carla	<i>¿Tienes frío?</i>	Ø expletive	Socioeconomic Status
44	Andrea	<i>Se coge dos o las manzanas que quiera</i>	Ø partitive	Socioeconomic Status
45	Carla	<i>Los niños están en la playa</i>	locative <i>en</i>	Personal Attractiveness
46	Andrea	<i>Pasamos todo el día al mercado</i>	locative <i>a</i>	Socioeconomic Status
47	Esteban	<i>Me gusta esta camisa, per las hay mejores</i>	Ø partitive	Personal Attractiveness
48	Andrea	<i>¿No quieres ir al cine conmigo?</i>	Ø expletive	Personal Attractiveness

# Stable variation or change in progress? A sociolinguistic analysis of *pa(ra)* in the Spanish of Venezuela

Stephen Fafulas, Manuel Díaz-Campos & Michael Gradoville  
University of Mississippi / Indiana University / Arizona State University

Sociolinguistic analyses of *pa(ra)* “for” have found that both linguistic and social factors play a significant role in speaker use of the reduced (*pa*’) and full (*para*) forms (e.g., Bentivoglio et al. 2005). However, no study to date has analyzed the extent to which production of these forms is most indicative of stable variation or a change in progress. The current study investigates 160 interviews conducted with speakers from Caracas, Venezuela and offers a comprehensive analysis of the variables of socioeconomic class, age, gender, and style in order to determine whether *pa(ra)* is best considered a stable variable or a change in progress in this speech community. Results indicate that the reduction of *para* lacks the hallmarks typical of linguistic change in progress.

## 1. Introduction

Mention of the reduction of *para* “for” to *pa*’ dates back to the early 20th century (e.g., Pietsch 1911; Hanssen 1913: 50; Henríquez Ureña 1940).<sup>1</sup> Examples of this process can be found in excerpts from the Linguistic Atlas of the Iberian Peninsula, collected in the 1930s. However, with a few notable exceptions (see literature review below), no thorough quantitative analysis of this phenomenon was carried out until the work of Bentivoglio, Guirado, and Suárez in 2005, almost a century after the initial observations made by dialectologists. In that study, Bentivoglio et al. demonstrated that the reduction from *para* to *pa*’ and the variable use of both forms within and across speaker groups was indicative of ‘structured heterogeneity’

---

1. See Hanssen (1913) for other relevant references on the historical development of *para* > *pa*’.

(Weinreich, Labov & Herzog 1968). Similar to other investigations conducted under the paradigm of variationist sociolinguistics, their study revealed that through detailed empirical observation the frequency of each form (*pa'* vs. *para*) occurring in a specific linguistic environment and by a particular group of speakers could be predicted. A number of subsequent empirical analyses on this topic (to be reviewed shortly) confirmed that *pa(ra)* variation was prominent across a wide range of social situations, geographic regions and linguistic contexts. Thus, while not as extensively studied as other phenomena in Hispanic sociolinguistics, the *pa(ra)* variable might be available as a heuristic for determining speaker age, education, socioeconomic status, formality of the discourse, and geographic origin, similar to that of the reduction and retention of syllable-final and word-final /s/ throughout the Spanish-speaking world (e.g., Cedergren 1973; File-Muriel 2007; Lafford 1986; Poplack 1980).

*Pa(ra)* alternation should add something to what we already know about the general patterns of language change as well, given that we are investigating its status in the community in relation to social factors such as age, gender, socioeconomic level, and style. While we know that variation is happening in all systems of the grammar, at all times, and that the alternation between competing variants is a necessary stage in language change, variation does not necessarily entail change, nor will every initiated change reach completion (D'Arcy 2013). Thus, determining the status of a given variable at the speaker and community level and interpreting the course of direction of a potential linguistic change are complex issues (see Kyzar this volume; Labov 1994) that are constantly being refined within sociolinguistics. On the one hand, we might find that a particular variable displays stable alternation throughout the speech community and among individuals. Similarly, it is possible that both the speech community and speakers change together. However, we might also find that a variable is stable throughout the speech community but changes throughout the course of a speaker's life cycle (i.e., age-grading) or that a speaker retains a particular speech pattern but that the community introduces a change which spreads slowly (i.e., generational change). It is in these last two examples that matters become even harder to disentangle. Furthermore, as Sankoff (2006) observes, some linguistic features, such as individual words or phonetic innovations with little impact on the rest of the sound system, are easier to learn and thus able to diffuse even among older speakers in the speech community as opposed to features that are more structurally complex.

A primary example of the complexity of disentangling language variation and change is evident in the variable production of /ʒ/ (voiced vs. devoiced) in River Plate Spanish, which has been under observation since the 1930s (Fontanella de Weinberg 1979). A complex pattern of devoicing whereby older, more educated speakers disfavored the innovative variant while younger less educated speakers favored its use led Wolf (1984) to predict that the change would not reach comple-

tion. However, Chang (2008), amongst others, has recently posited that devoicing has steadily advanced and reached completion among certain social groups (see Rohena-Madrado 2011 for further discussion). In another example, Holmquist (1985) found that in rural Cantabrian speech younger speakers who gravitated toward urban life were likely to abandon word-final production of *-u* (i.e., *hablu* vs. *hablo*) given its association with village life, leading him to predict the eradication of this variant by future generations. Yet other variables, such as retention vs. deletion of intervocalic /d/ (i.e., *cantado* vs. *cantao*; see Lapesa 1981) have been in competition for centuries and appear to display a pattern of continuous, stable variation. In the case of *pa(ra)* alternation, Díaz-Campos, Fafulas, and Gradoville (2011) conducted multiple GoldVarb analyses and found that, while speakers from lower socioeconomic levels favored the reduced form overall, when age was entered into the analysis, the older speakers of the lower socioeconomic class as well as the younger speakers in the upper socioeconomic class favored reduction. Further, the middle socioeconomic group did not demonstrate a significant difference for the use of either form based on age group. Thus, more detailed analyses are necessary in order to determine the status of *pa(ra)* as a stable variable or one indicative of a change in progress. In order to achieve this goal, the present investigation expands on previous research and offers a more fine-grained analysis of the sociolinguistic variables of socioeconomic class, age, gender, and style in order to determine the status of *pa(ra)* variation in Caracas, Venezuela. This paper contributes to the Hispanic sociolinguistic and general sociolinguistic literature by offering a more complete analysis of the social variables involved in *pa(ra)* alternation and in determining how these variables can be used to interpret the status of this linguistic change in Caraqueño Spanish. Our analysis adds to other variationist studies which employ the apparent time construct in determining the direction of language change. We conclude by showing that only with a careful observation of an array of social factors is it possible to understand the status of *pa(ra)* alternation in this variety of Spanish.

## 2. Background literature

### 2.1 Determining patterns of stable variation vs. language change

Labov (2001b) reminds us that regular social and stylistic stratification are two key indicators of stable sociolinguistic variables. If all social groups in a given speech community are differentiated by their distribution of a variable but display a similar pattern of style shifting, this is indicative of stable variation. Chambers (2002) also highlights age as a primary social correlate of language variation and change: if a linguistic variant represents language change, it is marked by increasing frequency

down the age scale. This is important in the identification of the status of the *pa(ra)* variable analyzed in the current study, as a favoring tendency among younger speakers might indicate a sign of language change. Labov (2001b) posited the Curvilinear Principle as the hallmark of language change in progress: change (from below) originates in a central social group (i.e., middle class) in the socioeconomic hierarchy. Labov (1966, 1972) also detailed the importance of careful vs. casual speech, a definition of style based on different tasks such as an interview, reading a text, reading a word list, and reading a list of minimal pairs for the study of linguistic variation and change. Across his studies, Labov employed distinct manners of data collection in order to document the stylistic continuum, which ranges from least to most attention paid to speech. Among his pioneer efforts to experiment with the study of style is the 1966 work concerning /r/-deletion in New York City department stores. It was in this study that Labov used for the first time the secret and fast survey as a way to obtain spontaneous speech. This method is based in part on the Observer's Paradox: finding the most effective way to observe language used by speakers as if they were in a situation in which they would use vernacular/informal language. This is assumed to represent the most natural, informal speech of a speaker (i.e., their vernacular/language acquired in pre-adolescence). Labov argues that it is necessary to obtain speech samples closest to the vernacular as these reveal the most systematic data for linguistic observation and the detection of the rules governing variation (Labov 1984). Schilling-Estes (2002) explains that style-shifting may be intentional, involving self-conscious use of features of which the speaker and audience are very aware, or they may be subconscious, involving features that people do not even realize they are using. Further, these shifts can be short-lived, as when a speaker involved in a sociolinguistic interview momentarily shifts into a more vernacular style. Through these studies, we see that style can be observed through multiple tasks and/or in the observation of a single sociolinguistic interview.

Bell (1984, 1997), perhaps in a more elaborate conceptualization of style, holds that style-shifting is mostly a product of the influence that an audience has on the speaker in a given interaction (see also Coupland 2007). While this was seen as a critique to Labov's 'attention paid to speech' model, Labov (2001a) has clarified that style-shifting is both the result of adaptation to different audiences as well as the result of speaker audio-monitoring. Labov suggests that within the interview speakers may shift while speaking to the same interlocutor as if they were actually speaking to different audiences. In the current study, we partially adopt Labov's understanding of style-shifting by observing intra-speaker variation, through stylistic differences within the interview, where the interlocutors and social situation are roughly constant. Subsequently, we compare this variation across social groups and age cohorts. This notion is dependent upon the assumption that

during initial stages of the interview participants will have a heightened degree of audio-monitoring, and as the interview proceeds their attention paid to speech will decrease. Our design is somewhat innovative as an approach to documenting style-shifting as we analyze various periods of each recording and test the hypothesis that less audio-monitoring occurs in tandem with time, as speakers become more involved in the topics of the interview, thereby revealing more of their vernacular in later portions of the interview.

Hispanists have used these aforementioned concepts in a number of studies to analyze whether linguistic variables are best considered as stable or changing. For example, Díaz-Campos & Geeslin (2011) examined the social categories of age, gender, and socioeconomic class to determine whether the copula contrast [*ser/estar* “to be” + adjective] in Venezuelan Spanish is stable or advancing. Their analysis revealed that speakers from lower socioeconomic classes and older speakers favor *estar* while gender evidences no effect. Thus, they did not find a typical curvilinear pattern indicative of change in progress in which middle class groups would have favored innovation while groups at the opposite ends would not. However, their analysis of age and socioeconomic class combined suggests that younger speakers from upper and lower socioeconomic backgrounds favor *estar*, signifying that a change in copula choice may be slowly advancing in this variety of Spanish. Their investigation shows us that it is necessary not only to analyze the behavior of each variable in isolation but also to examine the interaction between variables in order to evaluate the status of a linguistic change. The current study employs this method to gauge the variation and change of the *pa(ra)* variable in a corpus of Venezuelan Spanish.

## 2.2 Background on *pa ~ para*

The alternation *pa'/para* is a widespread phenomenon, noted by several linguists of the 20th century (Bentivoglio et al. 2005), suggesting that these forms have most likely existed since the colonization of the Americas. However, documentation of the reduction to *pa'* is not extensively treated in the historical literature on Spanish (e.g., Lapesa 1981; Penny 1991; Zamora Vicente 1970). Recently, Gradoville (2014: 7) documented some of the disagreement amongst historical linguists regarding the origins of *para*. While authors such as Lathrop (1996) and Menéndez Pidal (1962) favored a *pro ad* etymology, Monlau (1881) considered *per ad* to be the original form. A third position, taken by Beardsley (1921), was that medieval Spanish *pōra* derived from *pro ad* and medieval Spanish *para* derived from *per ad*, with *para* subsequently entirely replacing *pōra*. Penny (1991), for his part, questioned the adequacy of explanation for the change to *para*. Gradoville (2014: 8–11) studied the rates of occurrence of different *para* variants in the *Corpus del Español* (Davies

2002–) and found that *póra/por a* and *para* were at roughly the same frequency of occurrence during the 1200s, the first century covered by the corpus. Thereafter, *póra* declined in frequency while *para* became the dominant form. With respect to the origin of *pa'*, Gradoville (2014: 10) reports infrequent occurrence of the form in early Spanish, which subsequently declined, possibly reflecting the standardization of orthography in favor of *para*. Hidalgo (2016) holds that reduction to *pa'* dates back to 1522, as evident in a letter by Hernán Cortés, and that the reduced form is still common in rapid speech throughout Latin America and Spain today, regardless of the socio-educational background of the interlocutors.

In an early empirical study on variable phenomena in Panamanian Spanish, Cedergren (1973) studied *pa(ra)* alternation and noted its relevance as a predictor of socioeconomic class and register formality. Cedergren (1973: 134) points out that the reduced form is prevalent in casual speech, but she also shows that reduction is socially stratified and not categorical in any particular group. Regarding age, Cedergren found that young adults (21–35 years of age) were less likely to use the reduced form of *para*. She also found that women favored the reduced form in Panamanian Spanish. García (1979) employed a language attitudes task and uncovered wide acceptability of the reduced form of *para* by speakers living in El Paso, Texas, and Ciudad Juárez, Chihuahua, with greater acceptability of *pa'* by the Southwest speakers. Hernández-Campoy & Jiménez-Cano (2003) studied radio broadcasts by Murcian Spanish speakers over a 26-year period. *Para* was among the set of linguistic variables examined in their study. Their results showed a change from 1975 to 1985 in favor of the full form *para*, which reached a production rate of 100% in the broadcasts of politicians in Murcia. Their study indicates the importance of age and formality in observing the *pa(ra)* variable. In their analysis, they proposed that radio broadcasts in Murcia are stylistically influenced by the norm found in the center and north of Spain, which these scholars interpreted as the national standard variety of language in Spain.

To our knowledge, the first study to systematically analyze the alternation between the full form *para* and the reduced form *pa'*, including both linguistic and extralinguistic factors, was Bentivoglio et al. (2005). Their data were from 48 sociolinguistic interviews extracted from the Corpus *Estudio Sociolingüístico de Caracas 1987* (Bentivoglio & Sedano 1993). Their participants were equally divided by age (30–45, 60+), socioeconomic level (upper, middle, lower) and gender (male, female). They analyzed the linguistic factors of following phonological context (vowel vs. consonant) and preposition meaning (purpose, directionality, temporality). The overall distribution of the 1,599 tokens analyzed revealed an even distribution such that 787 (49%) were of the full form (*para*) while 812 (51%) were of the reduced form (*pa'*). Results of the multivariate analysis conducted with GoldVarb showed that speakers from lower socioeconomic levels and the



60+ age group favored the reduced form, as did a following consonant (.54) and a directionality meaning (.66). Guirado (2007) analyzed the same linguistic and extralinguistic factors as Bentivoglio et al. (2005) but with the additional factor of dialect. The 72 speakers included in her study were from three distinct regions of Venezuela: Caracas, Mérida and Maracaibo. Similar to Bentivoglio et al. (2005), Guirado's analysis revealed that the use of *pa* and *para* was conditioned by social and linguistic factors. Of the 2,144 tokens analyzed in her study, 1,026 (48%) were of the reduced form *pa'* and 1,118 (52%) of the full form *para*. The patterning of her results was almost identical to that of Bentivoglio et al. (2005) with the addition that Mérida and Caracas favored the reduced form more so than Maracaibo.

Díaz-Campos et al. (2011) confirmed that social class stratification is primary in studying the *pa(ra)* variable, and they were able to show that increased retention (i.e., the full form, *para*) by younger generations was correlated with increased access to education in Venezuelan society. In a subsequent investigation, Díaz-Campos, Fafulas, and Gradoville (2012) employed a usage-based analysis of *pa(ra)* variation and illustrated that frequently used two-word combinations involving *para* (both preceding and following the preposition) were more likely to reduce than less frequent ones. Barrios (2012) analyzed perception of phonological processes, such as apocope in words including *para*>*pa'*, and found that native speakers of Uruguayan Spanish from Montevideo had a negative evaluation of these forms in the community. Bedinghaus (2013) analyzed a corpus of speech from Málaga, Spain, and found that when *para* appeared with frequently occurring function words it was significantly reduced, as it was in general among the older generation (55 years or more) and female speakers. These previous studies point to the fact that a variety of social and linguistic factors constrain use of *pa(ra)* in Spanish. However, no study to date has determined the extent to which this variable is best classified as a change in progress or one more characteristic of stable variation. The current investigation observes a set of sociolinguistic variables, including style, in order to decide whether *pa(ra)* is a change in progress or stable in the Caraqueño speech community.

### 2.3 Questions of the current investigation

The current study was guided by the following research questions:

- (1) Can *pa(ra)* variation be considered a stable variable? Or does the sociolinguistic patterning indicate a change in progress?
  - a. Does social class show a curvilinear pattern?
  - b. Do certain social groups show more stylistic variation than others?
  - c. Does reduction of *para* increase in the speech of the youth?
  - d. Do females use the innovative variant more than males?



With the relevant background literature and research questions detailed, the next section continues with a description of the methodology used in the present investigation.

### 3. Methodology

In the current study, we analyze 160 semi-directed Labovian-style sociolinguistic interviews from the corpus *Estudio sociolingüístico de Caracas* (Bentivoglio & Sedano 1993). Each interview consists of approximately 30 minutes of unplanned speech. As indicated by the procedures reported in Bentivoglio & Sedano (1993), two native-speaking Spanish speakers from Caracas, Venezuela conducted each recording. Inclusion criteria stipulated that all participants, and their parents, were born and raised in Caracas. All interviews were conducted in 1987–1988. This data is representative of the sociolinguistic profile of Venezuelan inhabitants in the eighties. Each interview period began with a question pertaining to the interviewee's childhood. Thereafter, the interviewers proceeded with particular topics of interest that were generated from the participants' accounts of their infancy or that came from a previously prepared list of topics, including the participant's occupation, educational background and religious beliefs as well as the political situation and typical holiday traditions of the region.

For the analysis, we exhaustively extracted all tokens of the dependent variable, full and reduced forms of *para*. The treatment of the variable in this analysis is categorical with two options, /para/ and /pa/. While we recognize the reality of intermediary productions, this method of conceptualizing the dependent variable as binary was selected as it is a traditional model of analysis employed in sociolinguistics. Within this study, unreduced *para* has been taken as the reference level, so all positive values in the statistical models are in the direction of reduced *pa*'. Each token was coded for each of the following variables.

#### 1. Socioeconomic level/class

The determination of socioeconomic level is based on a formula by Contasti (1980) and implemented by the researchers who designed the original corpus. The 160 participants were assigned to one of the following five groups: upper, upper-middle, middle, lower-middle, and lower socioeconomic level. Since Bentivoglio et al. (2005), Cedergren (1973), Díaz-Campos et al. (2011), and Guirado (2007) have all found the lower socioeconomic level to exhibit the most distinct speech pattern with respect to this variable, the lower socioeconomic level has been selected as the reference level and coefficients for the other socioeconomic levels are with respect to the lower socioeconomic level.

## 2. Age group

Speaker age refers to the age of the participant at the time of the interview. The designers of the corpus divided age groups as follows: 14–29, 30–45, 46–60 and 61 and older. Speaker age group has been treated as a continuous variable in this study due to the findings of, especially, Díaz-Campos et al. (2011) that the age trends for this variable were basically linear, at least within each social class. Each age group was assigned an integer: 1 (14–29 years), 2 (30–45 years), 3 (46–60 years) and 4 (61 years and older). In accordance with the recommendations of Gelman & Hill (2007: 55), for variables where the value zero is not meaningful in models with interactions, these values were centered (the mean of 2.5 was subtracted from these integers).

## 3. Gender

In total, the corpus is composed of 80 male and 80 female speakers. Female speakers were selected as the reference level, so any coefficients express how male speakers differ from the female speakers.

## 4. Style

In order to account for the effect of the trend suggested by Labov (1972) that a greater degree of attention is paid to speech at the beginning of an interview than at the end, a ratio was recorded to determine approximately how far the occurrence of *pa(ra)* was from the beginning or end of the interview. Each token was sequentially numbered within the interview. This value was then divided by the total number of tokens that the speaker produced in the interview, yielding a number between zero and one. Consequently, tokens with values closer to zero are relatively early in the interview while tokens with values closer to one are relatively late in the interview.

Statistical analysis was carried out using generalized linear mixed-effects models in the R programming language (R Core Team 2014). Mixed-effects models allow for the inclusion of fixed effects (socioeconomic class, age, gender) and random effects (individual). Using this type of model is a step beyond traditional logistic regression in that it addresses the issue of independence of observations, which is not directly accounted for in the implementation of a logistic regression model. Thus, random effects are useful for measuring the variability between subjects and the correlation of observations within subjects. Generalized linear mixed-effects models were identified as the best candidate for the analysis of the *pa(ra)* variable for the following three reasons. First, the findings of Díaz-Campos et al. (2011) demonstrated that different socioeconomic levels in Caraqueño Spanish do not behave the same way with respect to the age variable. Consequently, there is an interaction between age and socioeconomic level. Mainstream statistical packages facilitate the modeling of interactions between predictor variables

by permitting the inclusion of interaction terms in a model. The inclusion of interaction terms allows us to better see both the independent effect of our predictor variables and the combined non-independent effect of two or more predictors. Second, methods of analysis implemented through programs like SPSS and R allow for the modeling of continuous predictors and, as a result, make possible the consideration of age and style as continuous. Age, by nature, is a continuous variable. While age can be, and quite often is, broken up into categories, the divisions of these categories are typically somewhat arbitrary. While the metadata of the corpus only permit that age be treated in groups, modeling age as scalar allows us to retain the sequential, ordered nature of the variable, which is impossible if age is treated as a categorical variable. Style is another variable that, in these data, is a continuous variable. Since style here refers to the approximate location of the token with respect to the beginning and end of the interview, there is no natural division to divide the interview. The ability to model style as continuous allows us to avoid the imposition of categorical boundaries (and potential corresponding loss of information) that might arbitrarily place one token in one category and the one that follows it into the next. Finally, three of the predictor variables under consideration (socioeconomic level, age, gender) pertain to the individual speakers in the sample and not the token itself. The inclusion of variable intercepts for individual speakers allows us to estimate the significance of the predictors on the basis of the speakers, not the tokens.

The final model was chosen by way of the following model selection procedure. An initial model was fitted that included all predictor variables and all possible interaction terms involving two predictor variables. One by one, new models were fitted discarding the non-significant interaction term or predictor variable from the previous model with the highest *p*-value until only significant predictors and interaction terms remained. This procedure allows us to determine which variables significantly predict the reduction of *para* in the Spanish of Caracas and, thus, to judge whether *para* reduction is progressing in the speech community under observation or whether the pattern is more indicative of stable variation. A detailed description of the results is presented in the next section.

#### 4. Results

The present analysis is based on 5,208 tokens. Table 1 shows the distribution of the dependent variable in the entire corpus (raw tokens and percentage). As can be seen, the non-reduced form *para* constituted 3,157 (60.6%) of the total tokens in the corpus, while the reduced form *pa'* accounted for 2,051 (39.4%) of the total tokens.

**Table 1.** Overall distribution of *pa(ra)* in corpus of Caraqueño Spanish (in 1987–1988)

Form	Tokens	Percent
Full form: <i>para</i>	3157	60.6%
Reduced form: <i>pa'</i>	2051	39.4%
<b>Total</b>	<b>5208</b>	<b>100%</b>

Moving on to the statistical analysis of the extralinguistic factors explaining the choice between unreduced *para* and reduced *pa'*, Table 2 presents the significant main effects and interaction terms that remain upon the completion of the model selection procedure (see the Methodology section for more details). The significant main effects include age, social class, and style. The final model also includes an interaction term for age and social class, which tells us in which of the social classes the effect of age is significantly different from the general age effect. Note that neither gender nor any interaction term considered aside from age/social class produced a statistically significant result at any point during the model selection procedure. As such, these results confirm the lack of importance of participant gender that has been found in previous studies of this variable in Venezuelan Spanish (Bentivoglio et al. 2005; Díaz-Campos et al. 2011; Guirado 2007).

Table 2 follows several conventions. The random effects, in this case a variable intercept for speaker, are displayed at the top with the variance accounted for by the intercepts. The regular fixed effects are displayed thereafter (i.e., social class, style, age, age\*social class). Each fixed effect has been listed in the first column of the fixed effects in reverse order of the average absolute value of the z-value for the predictor. In the case of categorical variables (social class and its interaction with age), the reference level (the default value of the categorical predictor) has been listed first followed by the values contrasting with it. For each fixed effect listed, there are five columns. The first is an estimate of each coefficient. Positive values can be interpreted to be in the direction of reduction to *pa'*, while negative values can be interpreted to be in the direction of unreduced *para*. The coefficients for categorical predictors are with respect to the reference level. Higher magnitude values are indicative of a stronger effect within the predictor. The second column contains the standard error for the estimate, which as a positive number gives us an idea of how confident we can be of the estimate. The lower this value is, relative to the estimate, the more confidence there is that the estimate is accurate. The third column is the z-value, which is the regression estimate divided by its standard error. The z-value gives us a standard number that we can use to interpret how confident we are that the estimate is accurate. Higher magnitude z-values correspond with a greater degree of confidence. For example, the z-value of the upper class is  $-7.037$  while that of the middle class is  $-5.245$ , which indicates that

we can have more confidence in the estimate for the upper class. The fourth column displays the  $p$ -value, through which we determine significance level. Very low  $p$ -values have been rendered using scientific notation. The last column shows a standard symbolic representation of the significance level for each estimate on the basis of the  $p$ -value. Wholly insignificant estimates receive no symbol, insignificant coefficients with a  $p$ -value less than 0.1 receive a period (.), significant coefficients at the 0.05 level receive one asterisk (\*), significant coefficients at the 0.01 level receive two asterisks (\*\*), and significant coefficients at the 0.001 level receive three asterisks (\*\*\*). The following paragraphs discuss the effects of each of the significant predictors.

**Table 2.** Generalized linear mixed-effects model of social predictors of the reduction of *para* in Caraqueño Spanish.

Random effects:					
Groups	Name	Variance	Std.Dev.		
<i>Speaker</i>	(Intercept)	1.858	1.363		
Fixed effects:					
	Estimate	Std. Error	z value	p	Sig.
(Intercept)	0.20524	0.26073	0.787	0.43118	
Social Class					
<i>Lower</i>	**REFERENCE LEVEL**			n/a	
<i>Lower Middle</i>	−1.41394	0.36331	−3.892	9.95e-05	***
<i>Middle</i>	−1.93816	0.36954	−5.245	1.57e-07	***
<i>Upper Middle</i>	−1.70268	0.36544	−4.659	3.17e-06	***
<i>Upper</i>	−2.64188	0.37544	−7.037	1.97e-12	***
Style	0.51279	0.12268	4.180	2.92e-05	***
Age	0.61411	0.22581	2.720	0.00654	**
Age*Social Class					
<i>Lower</i>	**REFERENCE LEVEL**			n/a	
<i>Lower Middle</i>	0.15815	0.32568	0.486	0.62725	
<i>Middle</i>	−0.33756	0.33124	−1.019	0.30816	
<i>Upper Middle</i>	−0.05724	0.32499	−0.176	0.86020	
<i>Upper</i>	−1.05991	0.33375	−3.176	0.00149	**

The first predictor showing significant effects in Table 2 is that of social level/class. As we can see in Table 2, all other social classes show significant

differences when compared to the lower class. Specifically, all other social classes have a negative estimate, meaning that all other social classes exhibit significantly less reduction to *pa'* than the lower social class. The upper class, predictably, exhibits the highest magnitude coefficient ( $-2.64188$ ), which tells us that the upper class' speech patterns are the most distinct from the lower class' patterns, while the lower-middle class shows the lowest magnitude coefficient ( $-1.41394$ ), which means that its use of the variable is the most similar to the lower class.

Style was the second predictor showing significant effects in Table 2. An examination of Table 2 shows us that style has a positive coefficient ( $0.51279$ ). As higher values for style correspond with points later in the interview, the positive correlation indicates that as the interview progresses the probability of occurrence of reduced *pa'* increases relative to the beginning of the interview. Given the lack of a significant interaction effect between style and the other variables, we can say that this effect appears to be true of all speakers, regardless of their sociodemographic characteristics.<sup>2</sup>

Age was the third predictor exhibiting significant effects in Table 2. As we can see in Table 2, age has a positive coefficient ( $0.61411$ ), which means that as speaker age increases reduction to *pa'* also increases. In other words, younger speakers are using the *pa'* variant less than speakers in the older generation. It should be noted that this pattern does not apply throughout the data set, but this particular issue will be addressed in the following paragraph that deals with the interaction term.

The interaction term between social class and age also exhibited a statistically significant effect. If we examine Table 2, we can see that the upper class has a statistically significant coefficient for the interaction. This interaction effect's negative slope ( $-1.05991$ ) indicates that the upper class' pattern for age is negative relative to the overall age pattern ( $0.61411$ ). None of the other social classes' age trends differ from the lower class' in any statistically significant way.

The plot in Figure 1 facilitates our understanding of this relationship in the data. Figure 1 contains points for each speaker in the data set plotted at the coordinates of his/her age group and reduction rate. The figure also contains individual regression lines that show the age-based trend for each social class.

The results displayed in Figure 1 confirm that there is more reduction among the older speakers in the corpus as well as the lower socioeconomic

---

2. Regarding lexicalized forms, such as "*pa'fuera*, *pa'ca*", our findings show no evidence that these forms pattern any differently than other tokens of *para* with respect to social variables.

levels. As we can see in Figure 1, all social classes except the upper class have a positive slope between age and reduction. It is the upper class that differs in that it has a negative slope, which indicates that the youngest speakers in the upper class exhibit more reduction to *pa'* than the older generation. The plot in Figure 1 also shows us that, while in the oldest generation there is a strong three-way social class distinction, in the youngest generation the behavior of the non-lower classes has largely converged to create a two-way opposition between the lower class and the other social classes. This pattern whereby the linguistic variable in question creates a division resulting in two radically different social classes is what Labov called “sharp stratification” (1964a: 170). In other words, we can see a major division between the lower socioeconomic group and all other socioeconomic groups.

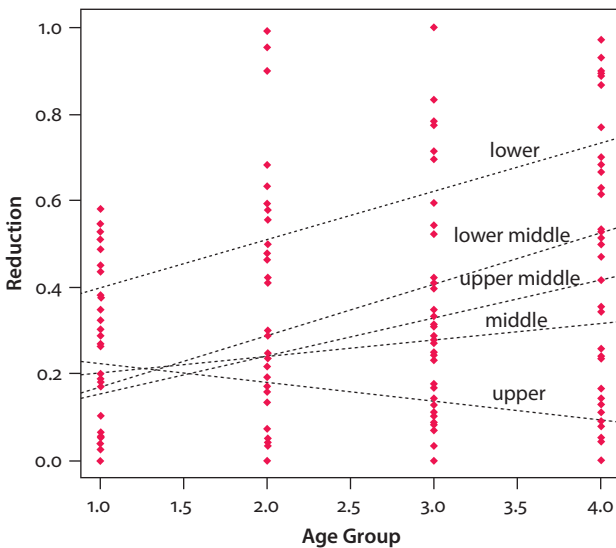


Figure 1. Plot of interaction between age and social class

Figure 2 depicts the interaction of participant age and style (or place in the interview). In Figure 2, age group is represented by different patterns and colors. The large thick solid black line is the overall trend between reduction and style/place in the interview. In Figure 2, social class is represented by line thickness. The upper socioeconomic class is represented by the thinnest line, while the lower socioeconomic class is depicted by the thickest line.

Of the twenty combinations of age and social class represented in Figure 2, only two (10%) of the combinations have a negative slope (upper, 46–60; lower-middle,

14–29)<sup>3</sup>, meaning that these two groups go against the overwhelming trend for speakers to use more *pa'* later in the interview. The fact that 90% of the data display a consistent trend lends credibility to the claim that *para* is sensitive to style, regardless of socioeconomic level and age. The thickest line at the top of Figure 2 represents the oldest group (61+) of lower class speakers, signifying that it is this group with the most overall reduction and use of *pa'*. We see that this group displays more reduction towards the end of the interview. The thin pink line is the oldest upper-class group. This group is at the bottom of Figure 2, meaning that their overall rate of reduction is the lowest amongst all groups and that they produced very little *pa'* at the beginning of the interview but increased *pa'* reduction towards the end of the interview.

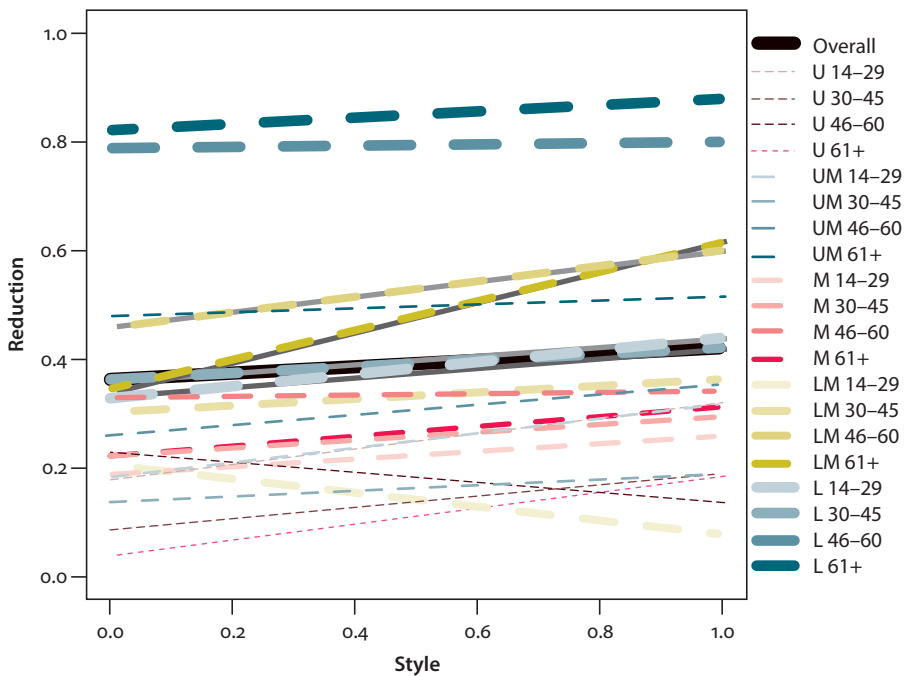


Figure 2. Plot of interaction between age and style

3. For more on the sociolinguistic backgrounds of these speakers and the general speech community, including educational trends, see Díaz-Campos, Fafulas & Gradoville (2011). This may help explain some of the findings reported here regarding the style trends of these speaker groups.



In the Discussion section, the implications of all factors, significant and not significant, will be observed with the ultimate goal of determining the status of the *para* variable in Caraqueño Spanish.

5. Discussion

Our observation of 160 speech samples from the Corpus *Estudio Sociolingüístico de Caracas* indicated a distribution of the *pa(ra)* variable dependent upon the factors of participant social class, age, and style/place within the interview. Equally as important, participant gender was not found to be a significant factor in determining use of *para*. In comparison to previous empirical analyses conducted on this variable in Venezuelan Spanish, our distribution can be seen in Table 3 below. It is important to note that a key difference between our study and those of previous authors is that in the current analysis we observe the behavior of all 160 speakers of the corpus, instead of a subset of the corpus as conducted by Bentivoglio et al. (2005) and Guirado (2007).

**Table 3.** *Pa(ra)* distribution compared to previous analyses of Venezuelan Spanish (Caraqueño, Merideño and Marabino).

Form	Bentivoglio et al. (2005)	Guirado (2007)	Current study
Full form <i>para</i>	787 (49%)	1,118 (52%)	3,157 (60.6%)
Reduced form <i>pa'</i>	812 (51%)	1,026 (48%)	2,051 (39.4%)
Total	1,599 (100%)	2,144 (100%)	5,208 (100%)

Table 3 reveals that as it relates to the overall distribution within the corpus of the current study there is a favoring of the full form *para*. However, importantly, the results of the statistical regression and corpus analysis showed that all socioeconomic groups, all age cohorts, and both male and female speakers, vary between these forms in their speech. In other words, both forms are prevalent in the speech community. We can conclude that the full and reduced variants constitute parts of the speech and/or grammars of the majority of individuals included in the corpus of Caraqueño Spanish.

The statistical regressions (Table 2) show us that socioeconomic level is an important variable in predicting participant use of the full or reduced form. All social classes displayed significant differences with the lower class, rendering negative estimates, thereby signifying less reduction to *pa'* than the lower social class. In fact, a trend was observed whereby the higher socioeconomic levels displayed a significantly greater use of the full form *para* than the lower socioeconomic

levels, revealing the following ordering of reduction to *pa'*: lower > lower-middle > upper-middle > middle > upper.

Figure 1 indicated that it was the lowest socioeconomic level which displayed the greatest favoring of the reduced form *pa'*. This effect was strongest among the older participants in the lower socioeconomic class group. The other socioeconomic groups showed a slight non-favoring effect or a disfavoring effect for their overall use of the reduced form *pa'*. Even within these groups, the pattern whereby older speakers typically favor more reduction holds. In other words, there is a pattern such that the higher the socioeconomic status of a participant group, the more likely they are to produce the full form *para*. The same is true of younger speakers who typically favor the full form in comparison to older speakers. This pattern whereby reduction decreases among the younger age cohorts might indicate that the use of *pa'* is declining in the speech community. That speakers in the higher socioeconomic groups are more selective in their use of the reduced form *pa'* could indicate that this form is considered a variant that is part of the vernacular. We can hypothesize that it may be the case that there is some consciousness of its vernacular nature in the Venezuelan speech community. For example, Hugo Chávez sought to appeal to the broader working class of Venezuela in his political campaigning and was often received by the public with the popular saying *Pa'lante comandante*. Díaz-Campos & Killam (2012) conducted a perception study in the Caraqueño speech community and found that deletion of syllable-final /ɾ/ was considered a vernacular variant typical of spontaneous speech. Sankoff & Wagner (2006), in their analysis of the inflected future in Montréal French, note that in successive stages of linguistic change variables typically transition from indicators, where social differentiation is not tied to stylistic differentiation; to markers, with both social and stylistic differentiation; or finally to stigmatization or stereotyping (see Labov 1972). We hold that the *pa'* form is a linguistic marker in the Caraqueño speech community. Through the analysis of the social profile of the variable, we can see its social stratification and predominant use by the lower classes and by speakers in informal/less audio-monitored styles. However, we note that additional analyses of language attitudes and perception of this variable in the speech community (e.g., Piqueres Gilabert & Fuss this volume) are necessary before any firm conclusions along these lines can be made.

Style was found to be a significant factor in determining speaker choice of the full and reduced forms of *para*. Labov (2001a: 86) summarizes the main findings of a variety of studies that have operationalized style: (i) For stable sociolinguistic variables, regular stratification is found for each contextual style; conversely, all groups shift along the same stylistic dimension in the same direction. (ii) In general, the range of social variation is greater than the range of stylistic variation. (iii) The second highest status group will generally show a greater style-shifting

slope than the others. (iv) Groups of speakers who are in contact with the speech community, but who are not an integral part of it and who thus can be considered outsiders, will often participate in the use of linguistic variables with altered stylistic patterns. (v) Stylistic variation is not found in the earliest stages of linguistic change but rather strengthens as the change progresses and is maximized if it is allotted prestige or social stigma. (vi) Style-shifting is related to the degree of social awareness of a linguistic variable by members of the speech community. The style analysis conducted for *pa(ra)* in the current study (see Table 1 and Figure 2) indicated a pattern whereby more reduction occurred as the interview proceeded, and this trend was consistent for the majority of age cohorts and socioeconomic levels. Further, as Labov suggests, our statistical analysis revealed that social level was significant. It seems that across the data, as participants began to become more comfortable during the interview and paid less attention to their speech, they revealed more of their vernacular and thus increased their use of *pa'*. This pattern correlates with our understanding of Labov's (1972) notion of the Observer's Paradox. Additionally, that there were no significant interactions between style and social class or age group signifies that all groups, regardless of their age or social status, follow this trend. We interpret these findings as an indication that speakers are somewhat conscious of the informal nature of the *pa'* variant, since it tends to be used in more relaxed contexts, and that this variant is stable and treated as such throughout the speech community. However, we also observed that the youngest lower-middle socioeconomic group, aged 14–29, showed a reversal of this pattern. Labov (2001a) also found that a group of younger speakers behaved contrary to the style trend for the variation of initial /θ/ and /ð/ consonants in words such as *the*, *this*, *then*, etc. In his earlier observations, Labov (1964b) noted that, although children begin to acquire the sociolinguistic norms of their particular speech community quite early, age and socioeconomic class constrain this process. Thus, the interaction of socioeconomic status, age, and style, particularly as they relate to the use of the *pa(ra)* variable, merits further analysis in a future study.

When taken together, we hold that our findings reveal that the *pa(ra)* variable is best considered a linguistic marker. Importantly, our assumption is based on the observation that *pa(ra)* varies according to socioeconomic class and style, thus indicating both cultural level and contextual style. The reduced form *pa'* is socially stratified by socioeconomic class and lacks the typical curvilinear pattern indicative of linguistic change in progress whereby the variant forms leading to a modification of the language are introduced and propagated by the intermediate groups. Thus, in our data, the overall reduction to the vernacular variant *pa'* lessens as socioeconomic status increases from lower to upper. Previous research holds that this pattern is most indicative of a stable variable (Labov 1994). Also, stylistic variation was evident within the sociolinguistic interview

itself, as the majority of socioeconomic groups used more *pa'* as the interview continued. Again, as we have already pointed out, Labov (2001a) notes that stylistic variation is not found in the earliest stages of linguistic change but rather strengthens as the change progresses and is maximized if it is allotted prestige or social stigma. While we do not hold that reduction to *pa'* is stigmatized in the Venezuelan speech community, it appears to be part of the vernacular and likely used more widely in less formal registers. Regarding age, all younger speakers, except for the upper socioeconomic class group, use less *pa'* than their older counterparts. Lastly, but importantly, is the fact that participant gender presented no noticeable effect, thereby adding to our claim that this variable is stable in the community under observation. In other words, no pattern in our data supports the general finding that women typically lead in linguistic change (Labov 2001b). Table 4 summarizes the observed patterns for the *pa(ra)* variable and compares whether each finding is more indicative of a change in progress or stable variation in the speech community.

**Table 4.** Sociolinguistic analysis of *pa(ra)* status in Caraqueño Spanish (adapted from Labov 1972, 1994, 2001a,b).

Factor	Change in progress	Stable sociolinguistic variation	Results for <i>pa(ra)</i> variable?
<i>Socioeconomic Level</i>	Curvilinear pattern	Non-curvilinear pattern	Non-curvilinear pattern
<i>Style</i>	More stylistic variation in certain groups	Consistent pattern for all social class groups	Consistent pattern
<i>Age</i>	Overall increased use of innovative form by youngest generations	No shift by younger groups to innovative form	Overall less shift by younger groups
<i>Gender</i>	Greater use of innovative form by female speakers	No differentiation by gender	No differentiation

Overall, our analysis indicates that *pa(ra)* can be considered a stable linguistic variable as it shows regular social stratification with more use of the reduced form by the lower socioeconomic class. Additionally, *pa(ra)* evidences regular stylistic stratification with more use of the reduced form towards the latter portion of the interview as speakers are presumably paying less attention to their speech and not monitoring. The general tendencies presented in Table 4 indicate that the *pa(ra)* variable has stabilized in the Caraqueño speech community. We recognize the need for a follow-up analysis with more recent data to confirm whether the variable has completely stabilized. This study joins a number of previous investigations which have used social factors and the apparent time construct (Bailey et al.

1991) as a means of observing linguistic variation and change (e.g., Silva-Corvalán 1986; Otheguy, Zentella, & Livert 2007), thereby adding to our understanding of the variable nature of *para* in Caraqueño Spanish.

## 6. Conclusions

The present analysis was based on 160 sociolinguistic interviews with native Spanish speakers from Caracas, Venezuela. The speakers in the corpus produced 5,208 total tokens of *para* and *pa'*. The non-reduced form, *para*, was found in 3,157 (60.6%) of the observed cases, and the reduced form, *pa'*, was found in 2,051 (39.4%). All socioeconomic groups showed variation in their speech with the lower socioeconomic group favoring the reduced form *pa'* significantly more than the middle and upper socioeconomic levels. In general, the analysis of stylistic variation, or attention paid to speech, indicated a pattern whereby the beginning of the interview showed less reduction to *pa'* than the middle and later parts of the interview. Regarding participant age, older speakers generally favored the reduced form *pa'*. No differences based on participant gender were found. Our sociolinguistic analysis indicated that the pattern of variation of *pa(ra)* is best considered stable since this linguistic variable showed regular social stratification as well as regular stylistic stratification with more use of the reduced form by the lower class and in the later, less-monitored portions of the interview. Thus, our study contributes to the findings of variationist sociolinguistics (e.g., Chambers 1995; Labov 2001b; Tagliamonte 2011) in revealing the intricate connection between the socio-cultural environment and language structure in Caraqueño Spanish. More specifically, the present investigation contributes to the Hispanic sociolinguistic and the general sociolinguistic literature with a complete analysis of the social variables involved in *pa(ra)* alternation, demonstrating how these variables, in correlation with the apparent time construct, can be used as a heuristic for interpreting the status of this change in Caraqueño Spanish. Further, our analysis of the distribution of *pa(ra)* as related to position within the interview, supports Labov's observation (2001a) that attention paid to speech (audio-monitoring) coupled with the presence of an interlocutor are both important factors in speaker style-shifting. Nevertheless, much work still remains on the use of the *pa(ra)* variable. For example, future research comparing the use and status of this form among different regions of the Spanish-speaking world would further our understanding of its importance as a sociolinguistic marker. Additionally, research into the status of this variable in other languages such as Brazilian Portuguese (see Gradoville 2014) will help us in our understanding of the evolution of this form in Romance languages. Moreover, a series of trend studies and panel studies (e.g., Sankoff &

Blondeau 2007) would allow us to confirm whether the status of the *pa(ra)* variable has changed in Caraqueño Spanish since the collection of the corpus used in the current study. Lastly, the historical trajectory of the reduction of *para* to *pa'* still remains open to future investigation.

## References

- Bailey, Guy, Tom Wikle, Jan Tillery, & Lori Sand. 1991. "The Apparent Time Construct". *Language Variation and Change* 3:3.241–264. doi:10.1017/S0954394500000569
- Barrios, Graciela. 2012. "Marcadores sociolingüísticos y prescripción idiomática". *Revista Digital de Políticas Lingüísticas* 4:4. Retrieved from [www.digilenguas.fl.unc.edu.ar](http://www.digilenguas.fl.unc.edu.ar).
- Beardsley, Wilfred A. 1921. *Infinitive Constructions in Old Spanish*. New York: Columbia University Press.
- Bedinghaus, Rob. 2013. "¿Vas *pa(ra)* Málaga? The Reduction of *para* in Málaga, Spain: Effects of Frequency, Syntactic Category, and Social Factors". *Selected Proceedings of the 16th Hispanic Linguistics Symposium* ed. by Jennifer Cabrelli Amaro, Gillian Lord, Ana de Prada Pérez & Jessi Elana Aaron, 238–252. Somerville, Mass.: Cascadilla Proceedings Project.
- Bell, Allan. 1984. "Language Style as Audience Design". *Language in Society* 13:2.145–204. doi:10.1017/S004740450001037X
- Bell, Allan. 1997. "Language Style as Audience Design". *Sociolinguistics: A Reader and Coursebook* ed. by Nikolas Coupland & Adam Jaworski, 240–250. Basingstoke: Macmillan. doi:10.1007/978-1-349-25582-5\_20
- Bentivoglio, Paola, Kristel Guirado, & Guillermina Suárez. 2005. "La variación entre *para* y *pa* en el habla de Caracas". *Estudios de Dialectología, Lexicografía, Lingüística General, Etnolingüística e Historia Cultural* ed. by Mariano Lozano Ramírez, 214–237. Bogotá: Instituto Caro Cuervo.
- Bentivoglio, Paola & Mercedes Sedano. 1993. "Investigación sociolingüística: Sus métodos aplicados a una experiencia venezolana". *Boletín de Lingüística* 8.3–35.
- Cedergren, Henrietta C.J. 1973. *The Interplay of Social and Linguistic Factors in Panama*. Ph.D. dissertation, Cornell University.
- Chambers, J.K. 1995. *Sociolinguistic Theory: Linguistic Variation and its Social Significance*. Oxford: Blackwell.
- Chambers, J.K. 2002. "Patterns of Variation Including Change". *The Handbook of Language Variation and Change* ed. by J.K. Chambers, Peter Trudgill & Natalie Schilling-Estes, 349–372. Malden, Mass.: Blackwell.
- Chang, Charles B. 2008. "Variation in Palatal Production in Buenos Aires Spanish". *Selected Proceedings of the 4th Workshop on Spanish Sociolinguistics* ed. by Maurice Westmoreland & Juan Antonio Thomas, 54–63. Somerville, Mass.: Cascadilla Proceedings Project.
- Contasti, Max. 1980. "Metodología para la medición del nivel socioeconómico para la población venezolana". *Boletín de la Asociación Venezolana de Psicología Social* 3:2.13–17.
- Coupland, Nikolas. 2007. *Style: Language Variation and Identity*. Cambridge: Cambridge University Press. doi:10.1017/CBO9780511755064
- D'Arcy, Alexandra. 2013. "Variation and Change". *The Oxford Handbook of Sociolinguistics* ed. by Robert Bayley, Richard Cameron & Ceil Lucas, 484–502. New York: Oxford University Press.

- Davies, Mark. 2002–. *Corpus del Español*. Retrieved from <http://www.corpusdelespanol.org>.
- Díaz-Campos, Manuel, Stephen Fafulas & Michael Gradoville. 2011. "Going Retro: An Analysis of the Interplay between Socioeconomic Class and Age in Caracas Spanish". *Selected Proceedings of the 5th Workshop on Spanish Sociolinguistics* ed. by Jim Michnowicz & Robin Dodsworth, 65–78. Somerville, Mass.: Cascadia.
- Díaz-Campos, Manuel, Stephen Fafulas & Michael Gradoville. 2012. "Variable Degrees of Constituency: Frequency Effects in the Alternation of *pa* vs. *para* in Spoken Discourse". *Selected Proceedings of the 14th Hispanic Linguistics Symposium* ed. by Manuel Díaz-Campos & Kimberly Geeslin, 75–87. Somerville, Mass.: Cascadia.
- Díaz-Campos, Manuel & Kimberly Geeslin. 2011. "Copula Use in the Spanish of Venezuela: Social and Linguistic Sources of Variation". *Spanish in Context* 8:1.73–94. doi:10.1075/sic.8.1.04dia
- Díaz-Campos, Manuel & Jason Killam. 2012. "Assessing Language Attitudes through a Matched-Guise Experiment: The Case of Consonantal Deletion in Venezuelan Spanish". *Hispania* 95:1.83–102.
- File-Muriel, Richard J. 2007. *The Role of Lexical Frequency and Phonetic Context in the Weakening of Syllable-Final Lexical /s/ in the Spanish of Barranquilla, Colombia*. Ph.D. dissertation, Indiana University.
- Fontanella de Weinberg & María Beatriz. 1979. *Dinámica social de un cambio lingüístico*. Mexico City: Universidad Nacional Autónoma de México.
- García, Maryellen. 1979. "*Pa(ra)* Usage in United States Spanish". *Hispania* 62:1.106–114. doi:10.2307/340085
- Gelman, Andrew & Jennifer Hill. 2007. *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge: Cambridge University Press.
- Gradoville, Michael. 2014. *A Comparative Usage-Based Approach to the Reduction of the Spanish and Portuguese Preposition para*. Ph.D. dissertation, Indiana University.
- Guirado, Kristel. 2007. "¿*Pa* donde va a agarrar?: La alternancia para-*pa* en tres comunidades de habla venezolana". *Interlingüística* 17.455–464.
- Hanssen, Federico. 1913. *Gramática histórica de la lengua castellana*. Halle: Max Niemeyer.
- Hernández Campoy, Juan Manuel, & José María Jiménez Cano. 2003. "Broadcasting Standardization: An Analysis of the Linguistic Normalization Process in Murcia". *Journal of Sociolinguistics* 7:3.321–347. doi:10.1111/1467-9481.00227
- Henríquez Ureña, Pedro. 1940. *El español en Santo Domingo*, vol. 5. Buenos Aires: Facultad de Filosofía y Letras de la Universidad de Buenos Aires, Instituto de Filología.
- Hidalgo, Margarita. 2016. *Diversification of Mexican Spanish. A Tridimensional Study in New World Sociolinguistics*. Berlin, Boston: De Gruyter Mouton. Retrieved 17 Oct. 2017, from <https://www.degruyter.com/view/product/473522>.
- Holmquist, Jonathan. 1985. "Social Correlates of a Linguistic Variable: A Study in a Spanish Village". *Language in Society* 14:2.192–203. doi:10.1017/S004740450001112X
- Labov, William. 1964a. "Phonological Correlates of Social Stratification". *American Anthropologist* 66:6.164–176. doi:10.1525/aa.1964.66.suppl\_3.02a00120
- Labov, William. 1964b. "Stages in the Acquisition of Standard English". *Social Dialects and Language Learning* ed. by Roger Shuy, Alva Davis & Robert Hogan, 77–104. Champaign: National Council of Teachers of English.
- Labov, William. 1966. *The Social Stratification of English in New York City*. Washington: Center for Applied Linguistics.
- Labov, William. 1972. *Sociolinguistic Patterns*. Philadelphia: University of Pennsylvania Press.



- Labov, William. 1984. "Field Methods of the Project on Linguistic Change and Variation". *Language in Use: Readings in Sociolinguistics* ed. by John Baugh & Joel Sherzer, 28–53. Englewood Cliffs, New Jersey: Prentice Hall.
- Labov, William. 1994. *Principles of Linguistic Change: Internal Factors*. Oxford: Blackwell.
- Labov, William. 2001a. "The Anatomy of Style-Shifting". *Style and Sociolinguistic Variation* ed. by Penelope Eckert & John R. Rickford, 85–108. Cambridge: Cambridge University Press.
- Labov, William. 2001b. *Principles of Linguistic Change: Social Factors*. Oxford: Blackwell.
- Lafford, Barbara. 1986. "Valor diagnóstico-social del uso de ciertas variantes de /s/ en el español de Cartagena, Colombia". *Estudios Sobre la Fonología del Español del Caribe* ed. by Rafael A. Nuñez-Cedeño, Iraset Paez Urdaneta & Jorge M. Guitart, 53–75. Caracas: La Casa de Bello.
- Lapesa, Rafael. 1981. *Historia de la lengua española*. Madrid: Gredos.
- Lathrop, Tom A. 1996. *The Evolution of Spanish*. 3rd ed. Newark, Del.: Juan de la Cuesta.
- Menéndez Pidal, Ramón. 1962. *Manual de gramática histórica española*. Madrid: Espasa-Calpe.
- Monlau, Pedro Felipe. 1881. *Diccionario etimológico de la lengua castellana*. 2nd ed. Madrid: Aribau.
- Otheguy, Ricardo, Ana Celia Zentella & David Livert. 2007. "Language and Dialect Contact in Spanish in New York: Toward the Formation of a Speech Community". *Language* 83:4.770–802. doi:10.1353/lan.2008.0019
- Penny, Ralph. 1991. *A History of the Spanish Language*. Cambridge: Cambridge University Press.
- Pietsch, Karl. 1911. "Zur Spanischen Grammatik (II)". *Zeitschrift für Romanische Philologie* 35.167–179. doi:10.1515/zrph.1911.35.2.167
- Poplack, Shana. 1980. "Deletion and Disambiguation in Puerto Rican Spanish". *Language* 56:2.371–385. doi:10.1353/lan.1980.0033
- R Core Team. 2014. *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Retrieved from <http://www.R-project.org>.
- Rohena-Madrado, Marcos. 2011. *Sociophonetic Variation in the Production and Perception of Obstruent Voicing in Buenos Aires Spanish*. Ph.D. dissertation, New York University.
- Sankoff, Gillian. 2006. "Age: Apparent Time and Real Time". *Encyclopedia of Language and Linguistics* ed. by Keith Brown, 110–116. Oxford: Elsevier. doi:10.1016/B0-08-044854-2/01479-6
- Sankoff, Gillian & Hélène Blondeau. 2007. "Language Change Across the Lifespan: /r/ in Montreal French". *Language* 83:3.560–588. doi:10.1353/lan.2007.0106
- Sankoff, Gillian & Suzanne E. Wagner. 2006. "Age-Grading in Retrograde Movement: The Inflected Future in Montréal French". *University of Pennsylvania Working Papers in Linguistics* 12:2.201–216.
- Schilling-Estes, Natalie. 2002. "Investigating Stylistic Variation". *The Handbook of Language, Variation and Change* ed. by J.K. Chambers, Peter Trudgill & Natalie Schilling-Estes, 375–401. Malden, Mass.: Blackwell.
- Silva-Corvalán, Carmen. 1986. "Bilingualism and Language Change: The Extension of *estar* in Los Angeles Spanish". *Language* 62:3.587–608. doi:10.1353/lan.1986.0023
- Tagliamonte, Sali A. 2011. *Variationist Sociolinguistics: Change, Observation, Interpretation*. Malden, Mass.: Wiley Blackwell.
- Weinreich, Uriel, William Labov & Marvin Herzog. 1968. "Empirical Foundations for a Theory of Language Change". *Directions for Historical Linguistics* ed. by Winfred Lehman & Yakov Malkiel, 97–195. Austin: University of Texas Press.
- Wolf, Clara. 1984. "Tiempo real y tiempo aparente en el estudio de una variación lingüística: Ensondecimiento y sonorización del yeísmo porteño". *Homenaje a Ana María Barrenechea* ed. by Lía Schwartz Lerner & Isaías Lerner, 175–196. Madrid: Editorial Castalia.
- Zamora Vicente, A. 1970. *Dialectología española*. Madrid: Gredos.





## PART IV

# Current issues in bilingual variation



## *El futuro es perifrástico*

### Future tense expression in a bilingual U.S. Mexican community

Kendall Kyzar

Louisiana State University

In this study of the expression of futurity in the Spanish of Mexicans residing in the Baton Rouge, Louisiana, metropolitan area, I explore the distribution of the variants of futurity – the morphological future (MF), the periphrastic future (PF) and the simple present (SP) – and the constraints conditioning their occurrence. The results indicate that the PF registers the highest rate of occurrence with a frequency of 72.0% to the detriment of the MF. In comparison with other studies (cf. Gutiérrez 1995; Orozco 2007a), the PF is the most favored variant at the expense of the others. The type of verb, as reported in previous studies (cf. Orozco 2005, 2007a), is the linguistic constraint that most strongly influences the expression of futurity. Regarding social constraints, level of education conditions the use of the future. The lack of statistical significance for gender suggests that women and men have similar sociolinguistic behaviors. This opens the possibility of exploring other linguistic variables in other bilingual Mexican communities to determine whether the social trends that are found are limited to the expression of futurity or to the communities under study.

#### 1. Introduction

##### 1.1 The history of Spanish in Louisiana

Since the early 16th century with the Spanish colonization of the New World and the creation of the Viceroyalty of New Spain, seated in Mexico City and encompassing an area which is now known as the Southwest, Spanish has been part of the linguistic landscape of the United States (Escobar 2010: 445). The presence of

Spanish, or ‘patrimonial Spanish’<sup>1</sup> as the Spanish spoken during this time is called (Moreno Fernández 1990: 201), could be found in territories that now belong to Florida, Louisiana, Texas, New Mexico, Arizona, California, Nevada, Utah, much of Colorado and even small parts of Wyoming, Kansas and Oklahoma (Escobar 2010: 445). With the arrival of the 17th and 18th centuries, there were numerous flourishing Spanish settlements in what is today the Southwest region of the United States: Santa Fe (New Mexico, in 1609), Mission (Arizona, in 1692), San Xavier del Bac (Arizona, in 1700) and San Antonio (Texas, in 1718). In 1803 Spain lost Louisiana (which first belonged to France and then to the United States), and in 1821 Florida was purchased by the new United States administration (Escobar 2010: 446).

As a European territory in the New World, Louisiana has a history that is both diverse and complicated (King 2011: 260). The Louisiana territory, unlike other regions of North America that were colonies of European nations, was under the control of France, Spain and Britain at different times throughout history. In 1763, after negotiations with Britain after the Seven Years’ War and more than half a century of French rule, Spain gained control of the region (Cummins 2008: 67). According to King (2011: 260),

[A]lthough Spain’s official control over the territory lasted for only 40 years, both the presence of Spanish governmental figures as well as the influx of Spanish speakers had significant and lasting effects on the linguistic and cultural landscape of North America.

In spite of the fact that Louisiana’s long association with French speakers has often caused modern researchers to overlook the influence of Spanish rule, there are studies on Louisiana Spanish. For example, authors such Armistead (1992, 2007), Din (1986, 1988), Lipski (1990, 2008) and Villeré (1972) have conducted extensive research on the *isleños*, the Spanish speakers from the Canary Islands who originally settled in Louisiana during the colonial period and who still live today in St. Bernard Parish, but in very small numbers. In more recent studies, Campos Molina (2009) examined several sociolinguistic characteristics of the Latino population of the metropolitan area of Baton Rouge, Louisiana, and found that Latinos in this geographic region believe their native Spanish has changed due to the influence of

---

1. The term ‘patrimonial Spanish’ refers to an archaic form of Spanish inherited from the colonists that arrived to present day United States between the 16th and 18th centuries. As explained by Moreno Fernandez (2007), patrimonial Spanish was formed from several varieties spoken in Spain and over time was passed down by the children of these original colonists and later from those that arrived from Mexico (Moreno Fernandez 2008: 201, 212).

other Latinos with whom they have constant contact and also by English influence. King (2011; this volume) analyzed business letters written in Spanish during the colonial era in Louisiana and found that governors of the territory repeatedly used the synthetic future, which derived from the Latin periphrasis *infinitivo + habere* “infinitive + have” and implied an obligation on the part of the addressee to carry out the indicated action (cf. Sedano 2007: 132). In addition, Vidal Covas (2013) compared the use of subject pronouns in the varieties of Puerto Rican Spanish spoken in Louisiana and Puerto Rico and found that overt pronominal usage rates in Louisiana and Puerto Rico (37% and 37.8% respectively) are consistent with what has been found for other Caribbean dialects (Lipski 1994: 241; López Morales 1992: 137). Given the few investigations of Louisiana Spanish, the results of the present study help fill a gap in Spanish sociolinguistics.

## 1.2 The Mexican community in the United States

Mexicans and Mexican-Americans are the largest Latino population that have shared the most extensive history with the United States (Lipski 2008: 75). Today, Mexicans still represent the largest number of Hispanic immigrants residing in the United States (Silva-Corvalán & Lynch 2009: 104). Compared to other Latino populations in the United States, Mexican-Americans are divided into three groups (Valdés 1988): (1) permanent immigrants, (2) temporary immigrants, who spend an average of ten to twelve weeks in the United States before returning to Mexico, and (3) cyclical immigrants, who are often migrant laborers (Lipski 2008: 78). When arriving in the United States, chain migration has been observed: they often settle in regions or neighborhoods where there are groups of individuals from the same place, which leads to the establishment of communities in which daily activities (e.g., industrial, commercial, social, domestic) can be carried out in Spanish, without the need to use English (Silva-Corvalán & Lynch 2009: 105). Such communities, according to Lipski (2008: 77), can be found in Los Angeles, Chicago, Houston, Phoenix, San Diego, Dallas, San Antonio, El Paso and San Jose (California).

## 1.3 The expression of futurity

In modern Castilian, the expression of futurity is a tripartite linguistic variable, that is, a variable that is made up of three forms to express the future tense: the periphrastic future (PF), the morphological future (MF) and the simple present (SP). Each of the forms is illustrated in example (1):

- (1) a. *Viajaré mañana.* (MF)  
 “[I] will travel tomorrow.”

- b. *Voy a viajar mañana.* (PF)  
 “[I] am going to travel tomorrow.”
- c. *Viajo mañana.* (SP)  
 “[I] am travelling tomorrow.”

Researchers such as Fleischman (1982), Bybee et al. (1991, 1994), Dahl (1985, 2000), Aaron (2006, 2007) and Sedano (2007) suggest that, since the appearance of the future tense in spoken language, an alternation has existed between these variants, mainly the PF and MF, causing instability between the future forms. According to Orozco (2005: 64), this “consequence of the so-called instability of futures ... stems from the tendency of future paradigms to be recast periodically from modal VPs.” A good example of this tendency is the modern use of the PF, which seems to be in direct competition with the morphological form when used to express a modal meaning (e.g., *No sé si ella vendrá a la fiesta* “I do not know if she will come to the party”).

In addition, investigations reveal that the periphrastic form tends to be the preferred way of expressing the future tense in modern spoken varieties of Spanish (cf. Gutiérrez 1995; Orozco 2005, 2007a; Lastra & Martín Butragueño 2010). Studies on this phenomenon posit that the distribution of the variants is not random but rather conditioned by their own semantics and certain extralinguistic factors (Claes & Ortiz-López 2011: 50). To address the differences in distribution between different Spanish-speaking countries and sociocultural levels, Bauhr (1992: 71) argues that the compound form *voy a cantar* “I am going to sing” is more widespread in America than in Spain and is heard more among children and in informal, illiterate speech and less in more careful, conscientious speech. In the same vein, Kany (1969: 192) notes that the preferred use of the periphrastic over the morphological form is commonplace but that in America its use has extended into domains beyond those of the norm in Spain. Westmoreland (1997: 388) suggests that this preference is largely due to the fact that this structure is easier to learn, especially in communities where there is a presence of bilingualism between Spanish and American English, since the periphrastic form also exists in the latter. Similarly, van Naerssen (1983) argues that the PF is easier to use than the MF, since it helps the speaker fill a pause while planning his or her thoughts that allows them to maintain the conversation.

According to Claes & Ortiz-López (2011: 51), studies analyzing the competitive alternation between the analytic and synthetic forms (Iuliano & De Stefano 1979; van Naerssen 1983; Silva-Corvalán & Terrell 1989; Sedano 1994; Hernández 1999; Becerra 2005; Orozco 2007a; Méndez-Vallejo 2008; Osborne 2008) have indicated that in modern Spanish, speakers prefer the analytic form to express futurity. From such works, we can draw two generalizations: (1) in spoken

language, the use of the PF greatly exceeds that of the MF and (2) the MF tends to indicate a modal meaning of doubt and conjecture with respect to a present event. Different variables have been identified that could influence a virtual alternation between these two forms. Among these can be found

*el estilo empleado, la inmediatez o alejamiento de la acción futura, la determinación o indeterminación del tiempo en que tendrá lugar dicha acción, así como su conexión o desconexión con el presente del hablante.*

“the style used, the immediacy or distance of the future action, the determinacy or indeterminacy of the time in which the action will take place, as well as its connection or disconnection with the presence of the speaker”. (Sedano 1994: 226)

When the SP is used as a variant of futurity, its use will always be temporal, meaning in “contexts in which the speaker is referring to a situation that, if it occurs, will take place in a moment [or some time] after the act of enunciation” (Gutiérrez 1995: 219, 222). In cases like these, the future tense is expressed by the presence of an adverbial specification or time reference that is implicit in the discourse (Gutiérrez 1995: 220). The ‘principle of distance’, a term coined by Silva-Corvalán (1990), suggests that speakers tend to choose the form that has the most temporal, spatial and psychological meaning when there are two or more forms that are in competition (Gutiérrez 1995: 222).

## 2. Previous studies

### 2.1 The expression of futurity in monolingual communities

In his study of Spanish spoken in Mexico City, Moreno de Alba (1977) seeks to determine the frequency of the MF. From 100 hours of conversations with 126 respondents, a total of 3,830 verb forms expressing futurity were collected, 374 of which were uses of the MF with a temporal value. Of greatest interest to the author is the context in which each variant appears most frequently. For example, the PF is used more in independent or main clauses and in subordinate noun, adjective and adnominal clauses, while the MF is used more in coordinated clauses and the SP in subordinate adverbial clauses. In conclusion, the investigator states that, in those statements not governed by verbs that require the infinitive or the subjunctive, the expression of the absolute future in the speech of participants can be summarized as follows: the MF with a frequency of 23.2%, the SP with 25.8% and the PF with 51%.

A common linguistic constraint that has been examined to determine alternative uses of the future forms is (in)definiteness. Based on data obtained by Silva-Corvalán & Terrell (1989) in their study of the expression of futurity in speech



samples from Puerto Rico, the Dominican Republic, Venezuela and Chile, the PF is the preferred form for the definite future, while the MF is favored in indefinite contexts. The researchers concluded that the participants were more likely to use the MF to express probability in the present or in the future. Based on this fact, they imply that the use of the MF in contexts of probability and doubt could

explicar las diferencias entre el uso levemente más elevado de la forma morfológica en casos de futuro indefinido y condiciones futuras que en los de futuro definido.

“explain the differences between the slightly more elevated use of the morphological form in cases of indefinite future and future conditions than in those of definite future” (201–207)

Méndez-Vallejo (2008) analyzes the temporal use (only those cases in which the two forms can be alternated) of the PF and the MF in the Spanish of Bogota, Colombia, in an oral corpus (Samper-Padilla et al. 1998) and a written corpus of readers' opinions of five online newspapers. Between the two data sets, a total of 352 tokens were obtained, 201 for the MF and 151 for the PF. Of the 144 samples taken from the oral corpus, 30 were the MF and 114 were the PF while, from the written corpus, a total of 208 samples were collected, of which 171 were the MF and 37 were the PF. Based on the distribution of the variants found in the oral data, the MF was used at a rate of 20.8%, while in the written data the MF registered a rate of 82.2%. Furthermore, for the PF, the percentages reveal that it was used with a frequency of 79.2% in the oral corpus and 17.8% in the written corpus. According to Méndez-Vallejo, these results were not expected since upper class speakers typically have a higher level of education. Based on the statistical analysis, the results indicate that there were four significant factors for the oral corpus (sex, number/person, verb class and proximity in time to the present moment) and five factors for the written corpus (sex, newspaper, verb class, use of negation and proximity).

In a more recent study of Mexican Spanish, Lastra & Martín Butragueño (2010) analyze the PF and the MF in the Spanish spoken in Mexico City. As a starting point and in accordance with Sedano (1994), they analyzed a total of eleven linguistic constraints and four social constraints. As is the case in previous studies (Silva Corvalán & Terrell 1989; Sedano 1994, 2007; Gutiérrez 1995; Orozco 2007a; Claes & Ortiz-López 2011), the researchers found that the rate of occurrence for the periphrastic form was significantly higher than that of the morphological form. To explain this distribution, the researchers posit that, although their informants still employ the morphological variant, its use is restricted to specific contexts and the cases in which it could alternate with the periphrastic form are very

few. As for the linguistic factors that had the greatest impact on the distribution of the PF, it was found that the possibility of alternation, the type of non-alternation, the grammatical person of the verbal tense, temporal distance (both explicit and interpreted), the epistemic characterization (both explicit and interpreted), the expressed intention and the inclusion of the verbal form in contexts with negation had the greatest conditioning effect. The most significant social constraints were level of education and age. Finally, the researchers suggest that the shift to the preferential use of the PF is a phenomenon that originates from below (“upward change” according to Penny 2000: 69) and is in its final stages.

In their study of the expression of futurity in the Spanish of San Juan, Puerto Rico, Claes & Ortiz-López (2011) investigate the linguistic factors related to the degree of certainty of the probability of a future action (grammatical person, temporal distance and the presence or absence of an indication of high certainty) and the social factors gender and age. The four forms analyzed were the MF, the PF, the SP and the continuous present (e.g., *Mañana a estas horas estoy trabajando* “Tomorrow at this time I will be working”) though the continuous present was eliminated as there was no occurrence of this variant with a future reference. They found that the PF was the variant used most often, followed by the SP and MF, respectively. The results indicate that first person motivates the use of the MF, while it disfavors the PF and the SP; second person supports the PF; third person rejects both the MF and the PF; and the sixth person (a group that combines (1) the expression of the second person plural *vosotros* and *ustedes*, (2) the third person plural *ellos/ellas* and (3) a person of an unspecified reference) opposes the MF and the SP and favors the PF. Regarding social variables, the authors found that men favor the SP and the MF and reject the PF, while women prefer the PF to the detriment of the SP; likewise, the younger the speaker, the lower the use of the PF and the greater the use of the SP.

## 2.2 Expression of futurity in bilingual communities

To my knowledge, the first empirical study of the expression of futurity in both monolingual and bilingual Spanish Mexican varieties was conducted by Gutiérrez (1995). The main objective was to compare and contrast the contexts of alternation between the PF, the SP and the MF and their modal and temporal uses in the bilingual variety of the Southwest region of the United States and the monolingual varieties of Morelia, Michoacán, and Mexico City. Regarding the distribution of the variants, the results show that for both speech varieties the PF has a higher frequency than the other forms, which appear to show a process of reduction. Moreover, the MF shows a lower frequency in the bilingual than in the monolingual variety of Mexico City.

Urrea & Gradoville (2006) research the syntactic and social factors that influence the use of the PF and MF in the Mexican variety of Spanish spoken in New Mexico and southern Colorado. The results indicate that the PF registered the higher rate of occurrence (76.5%) while the MF registered a much lower rate (23.5%). According to the results of the statistical analysis, there were only three factors that were statistically significant: meaning (futurity or modality), the presence of a verbal specification (such as the adverbial *mañana*) and age of the speaker. The results indicate that the MF is used more with the verbs *decir* “to say”, *ser* “to be” and *tener* “to have”, which make up half of the tokens for this variant. The authors conclude that the statistical significance of age points to a change in progress that indicates that the MF is reaching the end of its life cycle in New Mexican Spanish.

Orozco (2007a) explores the expression of futurity of Colombians in New York City and their counterparts in Barranquilla, Colombia. A total of nine linguistic constraints were analyzed: temporal distance (type of future), clause length, type of clause, grammatical number of the subject, grammatical person and animacy of the subject, adverbial specification, verb transitivity and regularity and length of inflection of the MF. Of these factors, the presence and position of the subject was the only one that did not reach statistical significance. For the distribution of the variants, the researcher found that the PF had the highest rate of occurrence, followed by the SP and the MF, results that are consistent with what has been found in other speech communities (cf. Ferrer & Sánchez 1991; Moreno de Alba 1970, 1977: 143; Zentella 1997: 190; Porcel 2005, Blas Arroyo 2008). Orozco found that the grammatical number of the subject reached statistical significance for the MF and the SP but not for the PF. The trends found in New York were the same for Barranquilla but at different levels.

As an extension of this study, Orozco (2007b) analyzes the social constraints on the expression of futurity in urban communities of Colombians in New York and Barranquilla, Colombia. For the informants from Barranquilla, there are three statistically significant social factor groups: sex of the speaker, conversation conditions (if the conversation was one-on-one with the interviewer, if a third party was present but did not participate or if a third party was present and did participate) and the complex group that combines the effects of social status and age of the speaker. The author reports that women favor the MF and men the PF (and for that reason the latter are considered ‘innovators’ since they are leading the progression of this change); when another speaker participates in the conversation, the use of the MF is favored; and that middle-class individuals born after 1960 are the only speakers who favor the MF.

For participants from New York City, Orozco found that gender, age, educational level and the complex group that combines the effects of age of arrival and length of residence in the United States were the four statistically significant

social factor groups. Just as with participants from Barranquilla, men from New York City also favor the SP. The results for the New York Colombians indicate that younger speakers use the MF at a lower frequency than the other variants and that older speakers favor the SP and disfavor the PF. Level of education was the only factor that reached statistical significance for all variants: individuals who have not completed high school strongly favor the MF while disfavoring the PF and the SP. On the other hand, individuals who have completed university studies in the United States tend to favor the PF at the expense of the other two variants. Finally, for the group combining age of arrival and length of residence in the United States, speakers who have lived more than 10 years in the United States favor the PF while those with a duration less than 10 years favor the MF and at the same time disfavor the PF.

### 2.3 Language contact

Since the coexistence of several languages within any society is pervasive, situations of language contact are no exception, but rather the norm (Gafarange 2007). Although in the majority of contexts of language contact a situation of bilingualism is implied – whether local, national or regional – bilingualism and language contact are two phenomena that represent different areas of inquiry in contemporary linguistics (Klee & Lynch 2009: 1). To explain alternation between Spanish and English in the United States, two different perspectives have been adopted: social and linguistic. From a social perspective, most of the research has focused on factors such as attitudes toward the phenomenon and its acceptance as a way to forge a multicultural identity (Rodríguez-González & Parafta-Couto 2012: 463). According to Pountain (2003: 254),

[T]he possibility that Spanish-English language alteration will provide the basis for a new, consistent language variety which will eventually itself form the basis of a standard or will remain the exclusive preserve of competent bilinguals will depend on factors such as education and a sense of community identity (which would tend to produce a more individual style of speech as a badge of the community).

In this sense, it can be concluded that, from a sociological point of view, alternating between Spanish and English reflects the context in which they are spoken and the attitude and self-identification of the speakers (Rodríguez-González & Parafta-Couto 2012: 464).

From a linguistic perspective, researchers such as Silva-Corvalán (1988, 1994a), Winford (2005) and McWhorter (2007), respectively, postulate that one of the processes that characterize linguistic contact is simplification. According to Ferguson (1982: 59), simplification includes a reduction in the inventory of

linguistic forms and functions and semantic extension, as well as the elimination of alternative structures. For example, Phillips (1967), Sánchez (1972) and Solé (1977) among others found that, in the Southwest region of the United States, the PF and the SP are used more to express futurity, an overextension that occurs at the expense of the MF (Klee & Lynch 2009: 237). Gutiérrez (1995), Silva-Corvalán & Terrell (1989) and Orozco (2005) note that the use of synthetic (morphological) forms is limited to modal uses (to express possibility or supposition) not only in bilingual speech but also in monolingual communities (Klee & Lynch 2009: 237). Given this fact, it should be no surprise that this form almost ceases to incorporate its temporal function in language contact situations in which formal language acquisition and use of Spanish are socially restricted (Klee & Lynch 2009: 237–238).

In his study of the bilingual region of Castellón, Spain, where Valencian and Castilian are in contact (see Piqueres Gilabert & Fuss, this volume), Blas Arroyo (2007: 280–281; 2008: 32) found that, although the PF is favored, because of this contact, the MF is retained at a relatively high level, since in Valencian the MF is used in certain syntactic contexts in which Castilian requires another verbal tense or mood. Gutiérrez (1995: 218–219) obtained similar results in his study of the bilingual Spanish variety spoken in various cities across the Southwest region of the United States and the monolingual varieties of educated Spanish in Mexico City and the popular Spanish of Morelia, Michoacán. He found that the PF was the most favored variant, but as in the study of Orozco (2007a) the situation of language contact accelerates the processes of change that are in progress in monolingual varieties.

### 3. Methodology

#### 3.1 Research questions

The main purpose of the present study was to analyze data obtained from a group of Mexican residents of Baton Rouge, Louisiana, in order to compare the impact of linguistic and social constraints on their use of the future tense. Based on the results of previous studies, the following five research questions served as a guide while interpreting the statistical results: (1) Which variant will have the highest/lowest rate of occurrence among the informants? (2) What is the pattern of occurrence of the three variants? (3) Which of the linguistic/social constraints will have the greatest impact on the expression of futurity? (4) Is there a variant that tends to be used more by women than men or vice versa? (5) Who are the linguistic innovators?

### 3.2 Geographic context: Baton Rouge, Louisiana

Based on data taken from the latest census (2010) of the United States, Escobar & Potowski (2015: 1, 10) report a total of 50,477,594 Hispanics in the United States while Latino populations in the Southwest represent 56.3 percent of this number. As for the state of Louisiana, according to the census figures presented in Table 1, the populations of the state, East Baton Rouge Parish and the city of Baton Rouge were 4,601,893, 444,526 and 230,058 inhabitants, respectively. People of Hispanic origin accounted for about 4.5%, 3.8% and 3.3% of the total population, respectively. Finally, considering the dissemination of the Mexican nationality throughout this geographic region, the figures indicate that there were 76,467 Mexicans residing in the state and 6,141 and 3,290 in East Baton Rouge Parish and the city of Baton Rouge, respectively. Since the census information presented here reflects only the documented residents, the figures could be higher if undocumented residents are taken into account.

**Table 1.** 2010 Census information (LA & BR)

	Louisiana	EBR (Parish)	BR (City)
Population	4,601,893	444,526	230,058
Hispanic origin	4.5%	3.8%	3.3%
Mexican	76,467	6,141	3,290

### 3.3 Participants

The data set explored in this study of Mexicans in Louisiana is made up of the *Corpus del Castellano Mexicano de Baton Rouge, Luisiana*, consisting of 7.44 hours of sociolinguistic interviews transcribed into 38,630 words or 96 pages. The conversations were recorded with ten participants, five men and five women, who reside in the metropolitan area of Baton Rouge, Louisiana. To gather demographic information, each informant was asked to complete a demographic questionnaire that solicited the following information: age, first name, gender, level of education, length of residence in Louisiana, previous residence in another state, socioeconomic status (upper/middle or lower class), linguistic repertoire (Spanish monolingual, Spanish dominant bilingual, English dominant bilingual or balanced bilingual), home language (if Spanish was spoken in the home) and linguistic generation (first, second or third; to avoid confusion, this question was explained to every participant). All of the participants share four demographic characteristics: all were born

in Mexico, are considered members of the working class, have lived in other states before settling in Louisiana and have completed elementary school. At the time of data collection, their ages ranged between 19 and 50 years old. Three were in their 20s, one was a teenager, two were in their 30s, three were in their 40s and one was 50 years of age. Five are monolingual Spanish speakers, one is considered bilingual with English dominance, three are bilingual with Spanish dominance and one is considered a balanced bilingual. Regarding level of education, one attended only one year of high school; five completed high school, one of whom wants to continue his studies; one finished high school and one semester of college; two graduated from college; and one is in her third year of college. Length of residence in Louisiana ranged from two to twenty-two years. Finally, according to the classification of sociolinguistic generations established by Escobar & Potowski (2015: 62), nine of the participants are from the first generation (G1) and the tenth is from the second generation (G2).

### 3.4 Constraints analyzed

I examined the effects of nine linguistic and five social constraints (see the appendix for a table of constraints). For the linguistic factors, I studied those used in Orozco (2007a) to analyze their impact on the expression of futurity in the Spanish of Colombians in New York City. As classified by Orozco (2015), I analyzed the effects of the internal constraints that operate at three different morphosyntactic levels. Three of the constraints operate at the clause level: temporal distance (type of future), the length of the clause and the type of clause. Two function at the subject level: the grammatical number and the grammatical person and animacy of the subject. The remaining four operate at the predicate level: the presence of a specific time marker (adverbial specification), the transitivity and regularity of the verb and the length of the MF inflection. For the social factors, I tested a total of five constraints: the speaker's age, sex, level of education, socioeconomic status/occupation and the conversation conditions. To determine the conditioning effects of all the constraints, I used GoldVarb, a computer program that allows a multivariate analysis with binary applications (Tagliamonte 2006: 217). With a total of 793 tokens, I conducted a series of multiple regression analyses for each variant of futurity. These regressions involved running each variant of the future tense discussed herein against a list of internal and external factors and later comparing the results obtained for each variant against the others. I present the distribution of variants followed by a discussion of the linguistic and social factors that significantly influenced the expression of futurity among the participants.



## 4. Results and discussion

### 4.1 Distribution of the variants

The distribution of the variants of futurity presented in Table 2 provides vivid answers to research questions (1) and (2) and, based on what was stated earlier, reflects what happens in the rest of the Hispanic world. To determine how the variants were distributed among the ten speakers, I divided the total number of tokens for each variant by the total number of tokens between the three variants and produced the following results: the PF registered a frequency of 72.0%, the SP 25.3% and the MF 2.7%. Based on these results, it became clearly evident that the periphrastic form was chosen unanimously by all informants at the expense of the MF. Also, the upward progression from the MF to the PF made it possible to easily identify the order of preference among the participants – first the PF, followed by the SP and in last place the MF. Furthermore, given that the MF is used only 21 times by the participants, it appears that the SP and the PF have become the most preferred forms for expressing futurity at the expense of the MF. The following two examples (2)–(3) taken from the corpus represent two cases in which the MF was used. The utterances were taken from two male participants, both appear in declarative statements and one is an example of the morphologically regular future tense while the other is an example of the morphologically irregular future tense.

**Table 2.** Distribution of the variants

Variant	N	%
PF	571	72.0%
SP	201	25.3%
MF	21	2.7%
<b>Total</b>	<b>793</b>	<b>100.0%</b>

- (2) *Entonces, hoy en día sí **encontrarás** porque hay mucha gente ...* [CHL]  
 “So today yes, **you will find** it because there are many people ...”
- (3) *y nada más o **habrá** personas sí hispanas que...* [JHL]  
 “and nothing more or yes, **there will be** Hispanic people that ...”

Compared to studies of the expression of futurity in other speech communities, the distribution was closely similar. In his study, Orozco (2007a) found that both the Colombian population in New York City and that of Barranquilla, Colombia, preferred the use of the PF over any other form. Silva-Corvalán



(1994a) obtained similar results in her study of a bilingual variety of Spanish spoken by Mexican-Americans in Los Angeles, California. According to Gutiérrez (1995), the frequency of use of the morphological form has declined in bilingual speakers of the Southwest region of the United States, and the periphrastic form threatens to become the only one used to locate actions in the future. Although he did not analyze the SP in his study of peninsular Spanish in the bilingual region of Castellón, Spain, Blas Arroyo (2008) found that MF is decreasing due to a preferential use of the PF. Claes & Ortiz-López (2011: 59), who examined the SP in their study of the expression of futurity in San Juan, Puerto Rico, found that the analytic variant is the most favored among the speakers of their study. They attribute this sociolinguistic behavior to the fact that the PF has more versatile semantics causing it to appear more frequently.

#### 4.2 Distribution of the variants according to the speaker

Based on the information in Table 3, the percentages were higher for the PF than for any of the other variants, making it the most preferred among all the participants. Given that the percentages do not show a higher occurrence of the PF among men and a lower one among women or vice versa, speaker's sex did not reach statistical significance for this variant. Moreover, the majority of the young speakers, especially Alejandro (64.6%), Álvaro (76.4%), César (68.1%), Adriana (77.8%) and Graciela (84.2%) reported higher uses of the PF than their older counterparts, suggesting they are maintaining a more consistent use of this variant. Also, it was interesting to see how two of the informants, César (68.1% for the PF, 24.2% for the SP and 7.7% for the MF) and Esteban (64.7% for the PF, 27.1% for the SP and 8.2% for the MF), yielded similar percentages for the three variants, which seems to suggest that their use of the three variants shows a more uniform distribution. In general, since the PF registered the highest percentages of all three variants, we can conclude that this variant is the most uniformly distributed of them all.

Of the occurrence rates between the three variants, the percentages for the SP fall within the middle. As for the sex of the speaker, the percentages for men and women were relatively close, with the exception of those with very low percentages of the SP whose variant of choice was almost exclusively the PF and those with higher percentages of the SP that show an almost null use of the MF. Considering these facts, it becomes clear why sex was not a significant factor for the SP. Overall, the use of the SP is not as widespread as the PF but, nevertheless, still shows an overall consistent use throughout the community.

**Table 3.** Distribution of the variants according to the speaker

Speaker* (Age)	PF		SP		MF	
	N	%	N	%	N	%
Alejandro (28)	82/127	64.6%	43/127	33.9%	2/127	1.6%
Álvaro (35)	81/106	76.4%	25/106	23.6%	0/106	0.0%
César (24)	62/91	68.1%	22/91	24.2%	7/91	7.7%
Esteban (37)	55/85	64.7%	23/85	27.1%	7/85	8.2%
Rubén (42)	62/70	88.6%	8/70	11.4%	0/70	0.0%
Adriana (27)	34/45	77.8%	10/45	22.2%	0/45	0.0%
Graciela (19)	64/76	84.2%	12/76	15.8%	0/76	0.0%
Bianca (44)	31/45	68.9%	13/45	28.9%	1/45	2.2%
Jessica (40)	43/73	58.9%	28/73	38.4%	2/73	2.7%
Catalina (50)	56/75	74.7%	17/75	22.7%	2/75	2.7%
<b>Total</b>	<b>571/793</b>		<b>201/793</b>		<b>21/793</b>	

\* All participant names introduced in this chapter are pseudonyms.

In general, the percentages for the MF were extremely low or, in some cases, as with nearly half (four) of the informants, nonexistent. There were two men (Rubén 42 and Álvaro 35) and two young women (Adriana 27 and Graciela 19) who did not use the MF. Of the remaining participants, three (Bianca 2.2%, Jessica 2.7% and Catalina 2.7%) had similar rates of use, two (César 7.7% and Esteban 8.2%) showed higher percentages, though with a low frequency, and one (Alejandro 1.6%) recorded the lowest percentage. Overall, the distribution of the MF is sporadic and sometimes absent and therefore struggles to maintain a consistent presence throughout this speech community.

### 4.3 Significant linguistic constraints

The statistical weights in Table 4 show the conditioning effects of the four linguistic constraints that reached statistical significance for the three variants. As can be seen, both length of MF inflection and type of future were significant for all three variants, whereas type of clause was only significant for the PF and the SP and verb regularity only for the MF. The ranges together with their aggregate values for each variant are presented to illustrate the general conditioning effects of each constraint. A space in the table with no statistical information means that this factor did not reach statistical significance and was therefore eliminated from the

2. All participant names introduced in this chapter are pseudonyms.

regression analysis. In general, the results show that an individual factor strongly favoring the PF disfavors the SP and vice versa.

**Table 4.** Significant linguistic constraints

	PF	N	SP	N	MF	N
<b>Factor</b>						
<i>Length of MF inflection</i>						
<i>Ser</i>	.86	45	.29	11	.27	2
<i>Ver</i>	.52	19				
<i>Dar</i>						
<i>Ir</i>	.10	25	.93	66		
Disyllabic	.58	351	.42	91	.49	9
Polysyllabic	.46	131	.48	33	.74	10
Range (187)	.76		.64		.47	
<i>Type of clause</i>						
Declarative	.55	344	.45	92	–	–
Conditional	.18	59	.86	75	–	–
Negative (D)	.81	92	.20	7	–	–
Negative (C)	.37	41	.69	25	–	–
Interrogative +	.65	32	.17	2	–	–
Interrogative –	.20	3			–	–
Range (132)	.63		.69			
<i>Type of future</i>						
Distant	.71	276	.26	24	.75	17
Near	.35	162	.70	113	.40	3
Unbounded	.36	110	.64	59	.24	1
Undetermined	–	–	.45	5		
Range (131)	.36		.44		.51	
<i>Verb regularity</i>						
Regular	–	–	–	–	.24	3
Irregular	–	–	–	–	.69	18
Range (45)					.45	

As reported in previous studies (Orozco 2005, 2007a), the type of verb (according to the syllabic length of the MF inflection) is the constraint that most strongly influences the expression of futurity in the speech of Colombians in Barranquilla

and New York City. This tendency can also be observed in Table 4 given that it recorded the highest aggregate range of all the constraints analyzed. Starting with the MF, the figures indicate that this variant is most favored (.74) when used with polysyllabic verbs. As for the PF and the SP, the results show that when a single factor clearly favors the PF, it generally disfavors the SP. For example, the monosyllabic verb *ser* “to be” is favored (.86) when used with the PF and is disfavored (.29), along with the verbs *ver* “to see” and *dar* “to give”, when used with the SP. Also, it can be noted that the PF favors (.58) the use of disyllabic verbs whereas the SP disfavors (.42) the use of these verbs. The following statements from the corpus in (4) and (5) illustrate these preferences:

- (4) *la carrera que ella escogió este va a ser muy solicitada ...* [OML]  
 “the career that she chose is going **to be** very in demand ...”
- (5) *Mientras la gente cambie, vamos a tener un nuevo futuro...* [CHL]  
 “As long as people change, we are going **to have** a new future ...”

When the verb *ir* “to go” is used, the PF is strongly disfavored (.10) while the SP is strongly favored (.93). Given that for the PF and the SP similar trends between all speakers were reported, it can be concluded that the conditioning effect of this constraint is well disseminated throughout this Mexican speech community. Furthermore, since the verb *ir* is used to create the analytic form of the PF and was used more with the SP as shown in (6) and (7) below, it was expected that it would register the lowest statistical weights for the PF.

- (6) *tú vas creando esa imagen de tu padre.* [MHL]  
 “you **go** on creating that image of your father.”
- (7) *las puertas se van cerrando día a día ...* [MHL]  
 “the doors **go** on closing day by day ...”

These results provide further evidence of the gradual grammaticalization of the verb *ir*, which has gradually transformed from a lexical to an auxiliary verb – similar to what happened to the verb *haber* (cf. Penny 2000: 50; 2008: 237). Moreover, the scarce occurrence of *ir* in the PF is analogous to what happens in other languages in which verbs equivalent to *ir* rarely appear as main verbs in similar periphrastic constructions (Orozco 2004: 136).

With respect to type of clause, it can be seen in Table 4 that this constraint was statistically significant for both the PF and the SP but not for the MF. For the PF, participants favor the use of this variant in declarative clauses with (.55) or without negation (.81) whereas, for the SP, they prefer conditional clauses with (.86) or without negation (.69), respectively. These tendencies are illustrated in (8) and (9) from the corpus.

- (8) *te están pidiendo esta cosa **no vas a saber**, yo ya me pasó, ...* [JHL]  
 “they are asking you for this thing **you are not going to know**, it already happened to me ...”
- (9) *si tienes dinero para comprar, lo **compras**.* [KML]  
 “if you have money to spend, **you will buy** it.”

As for interrogative clauses, those without negation (.65) were statistically significant only for the PF. Furthermore, the results suggest that, in relation to the type of clause in which these two variants appear, particularly conditional and negative statements, the PF and the SP seem to be at opposite ends of the spectrum. Similarly, in the speech of Puerto Ricans and Colombians in New York City, Orozco (2015) found that, for the PF, both communities prefer negative statements (.58) while, for the SP, both groups favor conditional clauses (.54).

With regard to the type of future, the figures in Table 4 indicate that this constraint registered more statistical significance for the SP than for the PF and the MF since there were two significant types of future for the SP and only one type of future for the PF and the MF. The results show that the distant future is favored by both the PF (.71) and the MF (.75), whereas for the SP the near (.70) and unbounded (no defined time limit) (.64) futures are favored. The following three statements in (10)–(12) taken from the corpus reflect the preferred uses of the PF and MF as distant future and the SP as near future.

- (10) *en setenta años otro partido no lo **va a hacer**, ni en seis años, ni en doce años, ni en veinte años lo **va a hacer**.* [CHL]  
 “in seventy years another party is not **going to do** it, nor in six years, nor in twelve years, nor in twenty years is it going to do it.”
- (11) *y más le digo y ya, ya será que **será** más adelante le digo.* [JHL]  
 “and more I tell you, and it will be what it **will be** much later I tell you.”
- (12) *no sabes si al ratito o mañana te **agarra** inmigración y ya no, no vas a poder venir ...* [JHL]  
 “you don’t know if in a minute or tomorrow immigration **will catch** you and you no longer are going to be able to come ...”

Claes & Ortiz-López (2011: 63) and Orozco (2009a: 102) found that in the Caribbean speech of Puerto Rico and the coastal region of Barranquilla, Colombia, the SP is used in expressions of a short temporal distance. According to Orozco (2009a: 102), the favorable use of the SP in statements that point to the near future imply a context of proximity and provide a connection between the paradigm of the present tense and actions near the present. As for the use of the PF, Orozco (2009a: 102) states that, although this variant registered a low weight of .46, a usage of 42% indicates that the residents of Barranquilla in his study, just as the Mexicans (.71) in this study, prefer to use the PF in statements indicating a distant future. As

with Orozco (2009a), Lastra & Martín Butragueño (2010: 157) also report that the PF is associated more with distant and vague expressions in the speech of Mexico City. On the other hand, Sedano (1994: 232; 2007: 138) found that in the speech of Caracas, Venezuela, the PF is used not only in contexts of immediate or proximal posterity but also in contexts designed to indicate that the action will take place in a future perceived as distant or vague, as seen in (10) above with the use of adverbial specifications *in seventy years*, *in six years* and *in twelve years* before and after the verb. As for the MF, Orozco (2009a: 100), Claes & Ortiz-López (2011: 63) and Blas Arroyo (2008: 10) also found that the use of this variant is promoted when employed in statements that express a distant time frame. The use of adverbial specifications such as *más adelante* “much later” in (11) above points to a broad temporal distance. Based on what was proposed by Fleischman (1982), Orozco (2009a: 102) affirms that

*el efecto favorable de los enunciados en el futuro distante sobre el futuro morfológico concuerda con la naturaleza semántica de este paradigma, el cual se asocia a un concepto de posteridad o subsecuencia. El hecho de que los hablantes de castellano han continuado utilizando el futuro morfológico para indicar la carencia de inminencia de acciones futuras indica más aún que el futuro morfológico no ha perdido totalmente su dominio semántico tradicional de posteridad a manos de la perifrasis.*

“the favorable effect of distant future statements on the morphological future coincides with the semantic nature of this paradigm, which is associated with the concept of posterity or subsequence. The fact that speakers of Castilian have continued using the morphological future to indicate the lack of imminent future actions further indicates that the morphological future has not completely lost its traditional semantic domain of posterity at the hands of the periphrasis.”

When comparing all these results with findings of the present study, it seems that the use of the PF does not concur with its traditional use of immediacy as noted in other speech communities and that the SP has taken on a dual function since it is used with both the near (.70) and unbounded (.64) futures. In other words, since the MF has basically disappeared from the speech of the Mexicans in the present study, it seems that the SP has adopted the traits of temporality from the PF and that the PF has assimilated to the temporal characteristics of the MF. Nevertheless, what is clear is that, for the type of future, the SP and the PF seem to be in direct opposition.

Finally, the results in Table 4 indicate that verb regularity only reached statistical significance for the MF. As can be seen, the MF was favored more with irregular (.69) than with regular (.24) verbs. If we consider the results obtained for the length of MF inflection, it becomes obvious why irregular verbs registered a greater statistical weight given that they typically have more than two syllables and are therefore polysyllabic. The use of the MF with an irregular verb can be seen in (13) below taken from the corpus.

- (13) *Si nunca lo intentas, nunca sabrás y luego lo podrías lograr.* [CHL]  
“If you never try it, **you will never know** and later you could achieve it.”

4.4 Significant social constraints

The information presented in Table 5 responds to the third research question. As can be seen, of the five social constraints that were analyzed, only one was statistically significant: level of education. Based on this evidence, this constraint seems to provide a better indication of the social effects on the expression of futurity among the participants of this study. Furthermore, since during the statistical analysis all social constraints were eliminated for the PF, it can be concluded that the use of this variant in the speech community under study is not conditioned externally but internally. As for the SP, the results in Table 5 indicate that those speakers who have not completed high school favor this variant (.62). For the MF, a probability of .80 indicates that those who have completed university studies tend to express this variant more than the others.

Table 5. Significant social factors

	PF	N	SP	N	MF	N
Factor						
Level of education						
≠ high school	–	–	.62	43	.47	2
= high school	–	–	.52	95	.41	9
college out US	–	–	.47	13	.35	1
college in US/MX	–	–	.36	10		
≤ college	–	–	.39	40	.80	9
Range (71)			.26		.45	

In terms of the MF, the results were predictable; since it is the more traditional and formal variant used in academic settings, it could be expected that those speakers who have completed advanced studies would use this form more. Nevertheless, given the fact that the SP was used by speakers of every level of education, it can be concluded that the use of this variant is evenly distributed throughout the speech community.

Although the speaker’s sex was not statistically significant for this study (see Fafulas et al., this volume), it is worth mentioning the usage of the MF by both men and women to put into perspective the results obtained for level of education. Of the 21 tokens of the MF used among all the speakers, a total of 16 were used by men and 5 by women. In large part, these tendencies can be explained in

the following ways: (1) because the men have a higher level of education, this suggests that they were exposed more to the MF since traditionally, as it is the form that is taught and used in the academic world and (2) because the women have a longer length of residence in Louisiana, they have had longer contact with English, which, would increase the use of the PF (since this form also exists in English) at the expense of the MF.

#### 4.5 The linguistic innovators

Given that half of the participants in Baton Rouge fall within a similar age range (19 to 35 years old), it is the younger speakers of this study who demonstrate an innovative linguistic behavior, which explains the preferential use of the PF and provides answers to research question four. Based on the information in Table 6, among these speakers the variant that records the highest rate of occurrence is the PF with an average percentage of 74.2%. According to the hypothesis of Guy (1990: 52), which argues that linguistic innovations are promoted by young speakers, it seems that the younger speakers listed in Table 6 are the linguistic innovators of the Mexican metropolitan community in the city of Baton Rouge. These results validate another hypothesis by Guy (2000: 18), which postulates that linguistic factors conditioning linguistic change and variation are consistent within the different segments of a speech community. Furthermore, since it was determined from the results of the demographic questionnaire that nine of the ten informants are considered monolingual or bilingual speakers with dominance in Spanish, the results of this analysis appear to be consistent with the hypothesis of Silva-Corvalán (1994a: 208), which states that changes in language contact situations are internally motivated because “they are in progress in the ‘model’ monolingual variety before intensive contact with another language occurs”. In other words, it can be hypothesized that the preferential use of the PF in the speech of the participants of this study had already begun in Mexico before they came to Baton Rouge.

**Table 6.** The linguistic innovators

	PF	SP	MF
Speaker (Age)	%	%	%
Graciela (19)	84.2%	15.8%	0.0%
César (24)	68.1%	24.2%	7.7%
Adriana (27)	77.8%	22.2%	0.0%
Alejandro (28)	64.6%	33.9%	1.6%
Álvaro (35)	76.4%	23.6%	0.0%
Average	74.2%	23.9%	1.9%



## 5. Conclusion

Given that the results indicate that, overall, both linguistic and social factors weakly condition the use of the MF, it becomes clearer why this variant is so sparsely used. It also appears that the PF and the SP have become the most preferred forms used to express futurity to the detriment of the MF. For the participants of this study, this is due in large part to the convergence between English and Spanish that has basically led to the extinction of the MF. As argued by Orozco (2015: 366), the morphological future ceases to function as a marker of futurity; the expression of futurity, then, becomes a binary linguistic variable expressed via the periphrastic future and the simple present. This suggests that bilingual Mexican speakers in Louisiana are at a more advanced stage of the progressive change toward the analytical variant.

Orozco (2007a) states that, as Vulgar Latin evolved into what are now the modern Romance languages, the expression of futurity evolved and continues to evolve in Castilian. Therefore, other studies that have examined the expression of the future throughout the Hispanic world, not only in bilingual communities but also in monolingual communities, report that the PF is the dominant variant (Orozco 2007a: 313). Based on studies done by Gutiérrez (1995) and Orozco (2007a) that indicate that language contact situations accelerate changes that are latent in the language of the minority group, in this case the change in progress of the preferential use of the PF over the MF, it stands to reason that this study provides evidence of the same progression of change between the speakers who migrated from Mexico. As a result, this progression has led to an almost extinct use of the MF among Mexicans living in Baton Rouge. If we consider the length of residency of the participants, the most obvious reason pointing to this rapid change is the fact that half of them have resided in the United States for at least 12 years. Therefore, due to longer exposure to English, they tend to use the PF more since in English the same compound construction exists.

Thus far, everything mentioned about the future tense points to a process of grammaticalization. Bybee et al. (1994: 4–5) explain that

we do not restrict our interest in grammaticalization to the transition between lexical and grammatical status, but rather recognize the same diachronic processes at work in a long chain of developments. Included are changes in lexical morphemes by which some few of them become more frequent and general in meaning, gradually shifting to grammatical status, and developing further after grammatical status has been attained. The events that occur during this process may be discussed under the rubrics of semantic, functional, grammatical, and phonological changes . . .

On the other hand, Orozco (2007a) affirms that illocutions resulting from grammaticalization (Fleischman 1982: 103; Givón 1971; Schwegler 1990) are characterized by morphosyntactic uniformity. In the case of the expression of futurity, this means that the PF has become the default variant used to express futurity in Spanish. Furthermore, Orozco sees this as a pattern of “cyclicity”, in which a language changes from primarily synthetic to predominantly analytical, then becoming synthetic once again (2007a: 324–325).

The results from this study as well as those obtained from other monolingual and bilingual Mexican communities indicate that the MF is on the brink of vanishing from Mexican speech. Since the type of future was the only significant linguistic constraint for the MF, it can be expected that as the expression of futurity and Castilian continue to evolve, the MF will no longer be conditioned by internal factors and will disappear, resulting in a binary linguistic variable. Moreover, due to language contact between Spanish and English in Louisiana, it can be expected that the MF will disappear at a faster rate in this community than in Mexico, as the distribution of the variants appears to suggest. Since the PF was not conditioned by any social factor, it can be concluded that the use of this form is already widely distributed between the participants, which has ultimately contributed to the advanced decline of the MF.

This study provides another piece to the puzzle of morphosyntactic variation in the expression of futurity that exists throughout the Hispanic world. It also provides a point of departure for the study of Mexicans living in the metropolitan area of Baton Rouge, Louisiana.

## Acknowledgements

The author would like to express sincere thanks to Dr. Rafael Orozco for imparting his vast knowledge of the subject and to Dr. Jeremy King for his guidance throughout the writing and editing process. I will be forever grateful for all your help.

## References

- Aaron, Jessi E. 2006. “*Variation and Change in Spanish Future Temporal Expression*.” Ph.D. dissertation, University of New Mexico.
- Aaron, Jessi E. 2007. “El futuro epistémico y la variación: Gramaticalización y expresión de la futuridad desde 1600.” *Moenia* 13.253–274.
- Armistead, Samuel G. 1992. *The Spanish Tradition in Louisiana*. Newark, Delaware: Juan de la Cuesta.
- Armistead, Samuel G. 2007. *La tradición hispano-canaria en Luisiana*. Las Palmas de Gran Canaria: Anroart.

- Bauhr, Gerhard. 1992. "Sobre el futuro cantaré y la forma compuesta voy a cantar en español moderno." *Moderna Sprak* 86:1.69–79.
- Becerra, Servio. 2005. "Uso del futuro en el español hablado en Bogotá: El uso coloquial y la norma culta." *Actas del XIV congreso internacional de la Asociación de Lingüística y Filología de América Latina (ALFAL)* ed. by Alba Valencia-Espinoza, n.p. Monterrey: ALFAL.
- Blas Arroyo, José L. 2007. "El contacto de lenguas como factor de retención en procesos de variación y cambio lingüístico." *Spanish in Context* 4:2.263–291.  
doi: 10.1075/sic.4.2.07bla
- Blas Arroyo, José L. 2008. "The Variable Expression of Future Tense in Peninsular Spanish: The Present (and Future) of Inflectional Forms in the Spanish Spoken in a Bilingual Region." *Language Variation and Change* 20:1.85–126.
- Bybee, Joan, William Pagliuca & Revere Perkins. 1991. "Back to the Future". *Approaches to Grammaticalization*, Vol. II ed. by Elizabeth Traugott & Bernd Heine, 17–58. Amsterdam: John Benjamins.
- Bybee, Joan, William Pagliuca & Revere Perkins. 1994. *The Evolution of Grammar: Tense, Aspect, and Modality in the Languages of the World*. Chicago: University of Chicago Press.
- Campos Molina, Dally. 2009. "Sociolinguistic Characteristics of the Latino Population in the Baton Rouge Metro Area." M.A. thesis, Louisiana State University.
- Claes, Jeroen & Luis A. Ortiz-López. 2011. "Restricciones pragmáticas y sociales en la expresión de futuridad en el español de Puerto Rico." *Spanish in Context* 8:1.50–72.  
doi: 10.1075/sic.8.1.03cla
- Cummins, Light T. 2008 [1994]. "Louisiana as a French Colony." *Louisiana: A History*, 5th Edition ed. by Bennett H. Wall, 53–68. Wheeling, Ill.: Harlan Davidson.
- Dahl, Östen. 1985. *Tense and Aspect Systems*. Oxford: Blackwell.
- Dahl, Östen. 2000. "The Grammar of Future Time Reference in European Languages." *Tense and Aspect in the Languages of Europe* ed. by Östen Dahl, 309–328. Berlin: Mouton de Gruyter.  
doi: 10.1515/9783110197099.2.309
- Din, Gilbert C. 1986. "The Canary Islander Settlements of Spanish Louisiana: An Overview." *Louisiana History: The Journal of the Louisiana Historical Association* 27:4.353–373.
- Din, Gilbert C. 1988. *The Canary Islanders of Louisiana*. Baton Rouge: Louisiana State University Press.
- Durán Urrea, Evelyn & Michael Gradoville. 2006. "Variation in the Future Tense of New Mexican Spanish." *SALSA – Symposium About Language and Society – Austin*. Retrieved from <http://studentorgs.utexas.edu/salsa/proceedings/2006.htm>.
- Escobar, Anna M. 2010. "El español en los Estados Unidos." *Introducción a la lingüística hispánica* ed. by José Ignacio Hualde, Antxon Olarrea, Anna M. Escobar & Catherine E. Travis, 391–444. Cambridge: Cambridge University Press.
- Escobar, Anna M. & Kim Potowski. 2015. *El Español de los Estados Unidos*. Cambridge: Cambridge University Press.
- Ferguson, Charles. 1982. "Simplified Registers and Linguistic Theory." *Exceptional Language and Linguistics* ed. by Loraine K. Olber & Lise Menn, 49–66. New York: Academic.
- Ferrer, María C. & Carmen Sánchez. 1991. "El verbo y su función en el discurso." *Variación lingüística en el español de Rosario* ed. by Nélida E. Donni, María C. Ferrer, Carmen Sánchez & Susana Boretti, 45–108. Rosario, Arg.: Universidad de Rosario.
- Fleischman, Suzanne. 1982. *The Future in Thought and Language: Diachronic Evidence from Romance*. New York: Cambridge University Press.

- Gafarange, Joseph. 2007. *Talk in Two Languages*. Houndmills, UK: Palgrave Macmillan.  
doi:10.1057/9780230593282
- Givón, Talmy. 1971. "Historical Syntax and Synchronic Morphology: An Archaeologist's Field Trip." *Papers from the Seventh Regional Meeting of the Chicago Linguistic Society, April* 16–18, 7:1.394–415.
- Gutiérrez, Manuel J. 1995. "On the Future of the Future Tense in the Spanish of the Southwest." *Spanish in Four Continents: Studies in Language Contact and Bilingualism* ed. by Carmen Silva-Corvalán, 214–226. Washington: Georgetown University Press.
- Guy, Gregory R. 1990. "The Sociolinguistic Types of Language Change." *Diachronica* 7.47–67.  
doi:10.1075/dia.7.1.04guy
- Guy, Gregory R. 2000. "A identidade lingüística de comunidade de fala: Paralelismo interdialecc-tal nos padrões de variação lingüística." *Organon* 14.17–32.
- Hernández, Carmen. 1999. "La expresión de la futuridad en el español de Puerto Rico." *Revista Oclip* 9–10.53–57.
- Iuliano, Rosalba & Luciana De Stefano. 1979. "Un análisis sociolingüístico del habla de Caracas: Los valores del futuro." *Boletín de la Academia Puertorriqueña de la Lengua Española* 7.101–110.
- Kany, Charles E. 1969. *Sintaxis hispanoamericana*. Madrid: Gredos.
- King, Jeremy. 2011. "Power and Indirectness in Business Correspondence: Petitions in Colonial Louisiana Spanish." *Journal of Politeness Research* 7:2.259–283. doi:10.1515/jplr.2011.013
- Klee, Carol A. & Andrew Lynch. 2009. *El español en contacto con otras lenguas*. Washington: Georgetown University Press.
- Kyzar, Kendall. 2012–2014. *Corpus del castellano mexicano de Baton Rouge, Luisiana*. Baton Rouge: Louisiana State University.
- Lastra, Yolanda & Pedro Martín Butragueño. 2010. "Futuro perifrástico y futuro morfológico en el corpus sociolingüístico de la ciudad de México." *Oralia: Análisis del Discurso Oral* 13.145–171.
- Lipski, John M. 1990. *The Language of the Isleños: Vestigial Spanish in Louisiana*. Baton Rouge: Louisiana State University Press.
- Lipski, John M. 1994. *Latin-American Spanish*. London & New York: Longman Linguistics Library.
- Lipski, John M. 2008. *Varieties of Spanish in the United States*. Washington: Georgetown University Press.
- López Morales, Humberto. 1992. *El español del Caribe*. Madrid: Editorial MAPFRE.
- McWhorter, John. 2007. *Language Interrupted: Signs of Non-Native Acquisition in Standard Language Grammars*. Oxford: Oxford University Press.  
doi:10.1093/acprof:oso/9780195309805.001.0001
- Méndez-Vallejo, Dunia C. 2008. "Periphrastic and Morphological Future Forms in Bogotá Spanish: A Preliminary Sociolinguistic Study of Upper Class Speakers." *IULC Working Papers Online* 8:3.1–22.
- Montes Giraldo, José J. 1962. *Sobre la categoría de futuro en el español de Colombia*. Bogotá: Instituto Caro y Cuervo.
- Moreno de Alba, José G. 1970. "La expresión verbal de lo futuro en el español hablado en México." Ph.D. dissertation, Universidad Nacional Autónoma de México.
- Morena de Alba, José G. 1977. "Vitalidad del futuro de indicativo en la norma culta del español hablado en México." *Estudios sobre el español hablado en las principales ciudades de América* ed. by Juan M. Lope Blanch, 129–146. Mexico City: Universidad Nacional Autónoma de México.

- Moreno Fernández, Francisco. 1990. *Metodología sociolingüística*. Madrid: Gredos.
- Moreno Fernández, Francisco. 2007. "Social Remarks on the History of Spanish." *International Journal of the Sociology of Language* 184.7–20.
- Moreno Fernández, Francisco. 2008. "Dialectología hispánica de los Estados Unidos." *Enciclopedia del Español en los Estados Unidos* ed./coord. by Humberto López Morales, 200–221. Madrid: Santillana/Instituto Cervantes.
- Orozco, Rafael. 2004. "A Sociolinguistic Study of Colombian Spanish in Colombia and New York City." Ph.D. dissertation, New York University.
- Orozco, Rafael. 2005. "Distribution of Future Time Forms in Northern Colombian Spanish." *Selected Proceedings of the 7th Hispanic Linguistics Symposium* ed. by David Eddington, 56–65. Somerville, Mass.: Cascadilla.
- Orozco, Rafael. 2007a. "The Impact of Linguistic Constraints on the Expression of Futurity in the Spanish of New York Colombians." *Spanish in Contact: Policy, Social, and Linguistic Inquiries* ed. by Richard Cameron & Kim Potowski, 311–328. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/impact.22.22oro
- Orozco, Rafael. 2007b. "Social Constraints on the Expression of Futurity in Spanish-Speaking Urban Communities." *Selected Proceedings of the 3rd Workshop on Spanish Sociolinguistics* ed. by Jonathan Holmquist, Augusto Lorenzino & Lotfi Sayahi, 103–112. Somerville, Mass.: Cascadilla.
- Orozco, Rafael. 2009a. "El castellano del Caribe colombiano a comienzos del siglo XXI." *Revista Internacional de Lingüística Iberoamericana* 7:2.95–113.
- Orozco, Rafael. 2015. "Castilian in New York City: What Can We Learn From the Future?" *New Perspectives on Hispanic Contact Linguistics in the Americas* ed. by Sandro Sessarego & Melvin González-Rivera, 347–371. Madrid & Frankfurt: Iberoamericana/Vervuert.
- Osborne, Samantha D. 2008. "Variable Future Tense Expression in Andalusian Spanish." M.A. thesis, University of Georgia.
- Penny, Ralph. 2000. *Variation and Change in Spanish*. New York: Cambridge University Press. doi:10.1017/CBO9781139164566
- Penny, Ralph. 2008. *Gramática histórica del español, segunda edición actualizada*. Barcelona: Ariel.
- Phillips, Robert. 1967. "Los Angeles Spanish: A Descriptive Analysis." Ph.D. dissertation, University of Wisconsin–Madison.
- Porcel, Jorge. 2005. "Distancia temporal vs. modalidad: Contraste en el futuro simple del indicativo del español hablado actual." *Lingüística Española Actual* 27:1.63–93.
- Pountain, Christopher. 2003. *Exploring the Spanish Language*. London: Arnold.
- Rodríguez-González, Eva & M. Carmen Parafita-Couto. 2012. "Calling for Interdisciplinary Approaches to the Study of Spanglish and its Linguistic Manifestations." *Hispania* 95:3.461–480.
- Samper-Padilla, José A., Clara E. Hernández Cabrera & Magnolia Troya Déniz, eds. 1998. *Macrocorpus de la norma lingüística culta de las principales ciudades del mundo hispánico*. Las Palmas: Universidad de Las Palmas de Gran Canaria.
- Sánchez, Rosaura. 1972. "Nuestra circunstancia lingüística." *El Grito: A Journal of Contemporary Mexican-American Thought* 6.45–74.
- Schwegerler, Armin. 1990. *Analyticity and Syntheticity: A Diachronic Perspective with Special References to Romance Languages*. New York: Mouton de Gruyter. doi:10.1515/9783110872927
- Sedano, Mercedes. 1994. "El futuro morfológico y la expresión *ir a + infinitivo* en el español hablado de Venezuela." *Verba* 21.225–240.

- Sedano, Mercedes. 2007. "Future Tense Expressions in Several Spanish Corpora." *Working with Spanish Corpora*, 132–144. New York: Continuum.
- Silva-Corvalán, Carmen. 1988. "Oral Narrative along the Spanish-English Bilingual Continuum." *On Spanish, Portuguese and Catalan Linguistics* ed. by John Staczek, 172–184. Washington: Georgetown University Press.
- Silva-Corvalán, Carmen. 1990. "Current Issues in Studies of Language Contact." *Hispania* 73.162–176. doi:10.2307/343006
- Silva-Corvalán, Carmen. 1994a. *Language Contact and Change*. Spanish in Los Angeles. Oxford: Oxford University Press.
- Silva-Corvalán, Carmen & Andrew Lynch. 2009. "La demografía hispánica en suelo norteamericano: Los mexicanos." *Enciclopedia del Español en los Estados Unidos* ed./coord. by Humberto López Morales, 104–111. Madrid: Santillana/Instituto Cervantes.
- Silva-Corvalán, Carmen & Tracy Terrell. 1989. "Notas sobre la expresión de futuridad en el español del Caribe." *Hispanic Linguistics* 2.191–208.
- Solé, Yolanda. 1977. "Continuidad/descontinuidad idiomática en el español tejano." *The Bilingual Review/La Revista Bilingüe* 4.189–199.
- Tagliamonte, Sali A. 2006. *Analysing Sociolinguistic Variation*. New York: Cambridge University Press. doi:10.1017/CBO9780511801624
- Valdés, Guadalupe. 1988. "The Language Situation of Mexican Americans." *Language Diversity: Problem or Resource?* ed. by Sandra McKay & Sau-Ling Cynthia Wong, 111–139. Cambridge & New York: Newbury House.
- van Naerssen, Margaret M. 1983. "Ignoring the Reality of the Future in Spanish." *Second Language Acquisition Studies* ed. by Kathleen M. Bailey, Michael H. Long & Sabrina Peck, 56–57. Rowley: Newbury House.
- Vidal Covas, Lee-Ann M. 2013. "El uso variable de los pronombres sujetos en el castellano puertorriqueño hablado en Luisiana y Puerto Rico." M.A. thesis, Louisiana State University.
- Villéré, Sidney L. 1972. *The Canary Islands Migration to Louisiana 1778–1783*. Baltimore: Genealogical Publishing.
- Westmoreland, Maurice. 1997. "The Dialectalization of Spanish Future Tense Usage." *Word* 48.3.375–395. doi:10.1080/00437956.1997.11432474
- Winford, Donald. 2005. "Contact-Induced Changes: Classification and Processes." *Diachronica* 22.2.373–427. doi:10.1075/dia.22.2.05win
- Zentella, Ana Celia. 1997. *Growing up Bilingual: Puerto Rican Children in New York*. Cambridge, Mass.: Blackwell.

# Appendix

## Linguistic and social constraints

### Linguistic constraints

<i>Clause level</i>	<i>Predicate level</i>
<i>Type of future</i>	<i>Presence of a specific time marker</i>
Immediate/near	Pre-verbal
Distant	Post-verbal
Unbounded	Both
Cannot tell	None
<i>Length of clause</i>	<i>Verb transitivity</i>
1–5 words	Intransitive
6–8 words	Transitive
9 or more words	Copulative
<i>Type of clause</i>	<i>Verb regularity</i>
Declarative	Regular
Conditional	Irregular
Interrogative	<i>Length of MF inflection</i>
Negative interrogative	<i>Ser</i>
Negative	<i>Dar</i>
Conditional negative	<i>Ir</i>
<i>Subject level</i>	<i>Ver</i>
<i>Grammatical number of subject</i>	Disyllabic
Singular	Polysyllabic
Plural	
<i>Grammatical person and animacy of subject</i>	
1st	
2nd	
3rd human	
3rd non-human	

---

**Social constraints**


---

***Age***

Teen/younger than 20

20–29

30–39

40–49

50–59

Older than 60

***Level of education***

Has not completed high school

Only completed high school

Attended college outside the US

Attended/is attending college in US/MX

Completed college or higher

***Conversation conditions******Sex***

One-on-one conversation

Male

Others present but did not participate

Female

Others present interacted with speaker

***Socio economic status/occupation***

Upper/middle class (white collar)

Lower class (blue collar)





## L2 Spanish in the U.S. and the question of motivation

### Changing trends in post-secondary language study

Terri Schroth & Bryant Smith

Aurora University / Nicholls State University

University students' motivation to study Spanish includes a number of practical or instrumental factors (e.g., ease in finding gainful employment and desire to travel) and intrinsic or integrative factors (e.g., personal growth and using Spanish to communicate with family and friends). To better understand students' current motivation to study Spanish, we examined data collected from university students of lower-division Spanish language courses from two different geographical regions in the United States. The participants completed questionnaires that sought to determine their motivation for studying Spanish. Concurrent with our hypothesis, results of this pilot study indicate that the students of Spanish who participated in this study express more practical (instrumental) motivations to learn the language. A desire to connect with one's heritage or to speak Spanish with family members proved to be a less important motivator than we had anticipated, as did the geographic region of the students.

#### 1. Introduction

Motivation plays an integral part of any university student's decision to enroll in a specific course, to declare a certain major or minor or even to decide which university to attend. Motivation to study a foreign language in college is particularly complex. A number of integrative and instrumental factors are often related to a student's perception of a language's importance in his/her community, connection to his/her heritage, future potential for economic gain, self-efficacy and a number of other factors (cf. Dörnyei 1994, 1998, 2009; Ely 1986; Hsieh 2008; McEwan & Minkle 1979; Negueruela-Azarola 2011; Ushioda 2009).

The present study attempts to determine and better understand the many factors that influence today's university students to enroll in Spanish courses, particularly in light of demographic changes and a competitive job market in many parts of the U.S. Based on previous research, we assumed that a desire to find employment would be a key factor in motivation to study Spanish, and we also believed that a desire to communicate in Spanish with one's family, friends or community members would also be among students' top priorities. This study attempts to answer the following main research question: What are the principal and current motivations (integrative and instrumental) for college students to study Spanish?

## 2. Previous research

Research on the topic of motivation in second language (L2) learning has seen numerous changes in specific subtopics and research variables, and it has experienced several time periods of renewed interest. Due to the large number of studies on motivation in L2 learning, this review highlights researchers who focus on integrative versus instrumental motivational factors and other areas that inform the present study, such as the strength of students' motivation to learn an L2. Pioneer researchers Gardner & Lambert (1972) examined the relationship between motivation and learners' attitude and L2 learning (see Piqueres Gilabert & Fuss, this volume, for insight into the concept of attitude in the context of the Spanish language). They provided a differentiation between the core of motivation, an orientation as an interpersonal quality (termed 'integrative orientation') and one that has a more practical quality ('instrumental orientation'). Learners who are motivated instrumentally learn a language for more extrinsic reasons, such as the possibility of a better job or reasons related to travel. In contrast, integratively motivated learners possess an internal desire and genuine interest to connect or integrate with the target language or culture, or they have an intrinsic goal, such as improving global mindedness. It is important to note that in previous studies researchers have used many different definitions and terms for the types of motivations. We use Gardner & Lambert's (1972) construct of integrative (desire to affiliate with the second language community or intrinsic personal goal) and instrumental (future career or practical uses) orientations to measure students' motivation for studying foreign language.

Gardner (1985) further defined motivation to learn another language in terms of the commitment that learners possess in studying a language and the desire they have to do so. Gardner (1985) added that this motivation could be assessed with the Attitude/Motivation Test Battery (AMTB) (Gardner et al. 1979), which became a standard instrument in measuring attitude/motivation. The AMTB

measures students' level of agreement ("strongly disagree" to "strongly agree") with certain statements regarding attitude, interest and motivation, and it contains a section on strength or rating of additional statements.

McEwan & Minkle (1979) asked 438 students of Spanish from California (200 total college students; others ranged from grades 3 to 12) 26 questions about five comparable areas of motivation to study Spanish (i.e., requirements or obligation; joy of learning/recreation; work/profession; global awareness/culture; and influence of friends or relatives). Results indicated that college learners of Spanish expressed high levels of motivation for the integrative areas such as the joy of learning and culture/global awareness and lower levels in instrumental motivation as a requirement/obligation (i.e., negative motivation, cf. McEwan & Minkle 1979) and work/profession. The percentages of college students who indicated the following motivations were as follows: 5% requirement (instrumental); 60% joy of learning (integrative); 40% work/self-improvement (instrumental/integrative); and 53% global awareness/cultural knowledge (integrative). McEwan & Minkle's (1979) research questions and instrument informed the present study, as their inventory of questions and answers they sought are similar to our inquiries. The top six most important reasons for studying Spanish at the college level were as follows: "it would be fun to learn and know a foreign language" (integrative); "desire to visit a foreign country" (instrumental); "enjoy reading in Spanish" (integrative); "it is useful for work" (instrumental); "it will help students find a better job (instrumental)"; and "it will help me understand other people better" (integrative).

Ely (1986) noted that a student may have both integrative and instrumental orientations and, perhaps, many different kinds of motivations. For example, the ability to speak with Spanish speakers in the U.S. could hold both integrative (internal desire to understand people of a culture) and instrumental (practical motivation to communicate) qualities. Therefore, an overlap of motivations may occur. In determining the type of motivation, Ely sought to find out more about these motivational factors and the strength of these motivations. Ely explored language learning motivation and developed a scale to measure the strength of motivation, using first-year university students studying Spanish. Ely tested his survey in pilot rounds to develop categories and clusters of motivation. Ely's pilot questionnaire gleaned 184 reasons for students taking Spanish, which he further organized into over 15 categories of instrumental and integrative motivations, such as use in study abroad, potential employment, university or department requirement and interest in learning other cultures. Ely's study (along with that of McEwan & Minkle 1979) became quite interesting to us and thus informed the present study.

The final version of Ely's survey was administered to a total of 75 first-year university students of Spanish from a university in northern California. This

round of surveys aimed to investigate the second goal of Ely's study: to discover the relationship between the type and strength of motivation. Factor analysis yielded three types of motivation clusters and three hypotheses about the factors: Cluster A – Integrative Motivation (communication/culture), positive influence; Cluster B – Instructional Motivation (job/marketability), positive influence; Cluster C – Requirement, statistically not significant/negative trend. Two clusters of reasons were found to contribute strongly to strength of motivation. Cluster A included a desire to communicate and to learn about another culture (i.e., integrative motivation). Cluster B included motivations of help with future career and ability to speak with Spanish speakers in the U.S. (i.e., instrumental motivation). Cluster C described taking Spanish because it was a requirement of some kind and "although requirement motivation was not a statistically significant source of variation in strength of motivation, the trend was a negative one" (Ely 1986: 32). A majority of participants reported a combined effect of both Clusters A and B. Ely emphasized, again, the complexity of student motivations, as types of motivation (integrative/instrumental) can overlap.

Dörnyei (1998: 117) discussed a marked renewal of interest in L2 motivation studies: "a number of researchers in various parts of the world attempted to reopen the research agenda in order to shed new light on the subject." The usage and discussion of the definition and scope of intrinsic vs. extrinsic motivation have been revisited by numerous researchers (cf. Deci 1992; Deci et al. 1991; Dörnyei 1998, 2009; Gardner & Clément 1990; Gardner & MacIntyre 1993; Paris & Turner 1994; Vallerand 1997). Dörnyei (1994) developed a framework to discuss L2 learner motivation self-system – the components of foreign language learning motivation. This framework separated L2 motivation into the three categories of language level, learner level and learning situation level. The language level involved integrative and instrumental goals or orientations to motivation on the part of the language learner. The learner level included traits specific to each learner, such as self-confidence, anxiety and perceived L2 competence. The learning situation level encompassed three motivational components: course-specific, teacher-specific and group specific. Teacher-specific and group-specific learning situations were external and could influence the motivation of the individual student or a group of learners. From this, Dörnyei's L2 motivational self-system model emerged and highlighted the importance of the L2 self as a powerful motivator.

Dörnyei's studies inspired many more recent investigations into such integrative and instrumental traits of the L2 self, including individual learner identity, individual differences, learner goals, learner interest, goal orientation, career and social aspiration motivation, among others (cf. Grenfell & Harris 2013; Iaccarino 2012; Kormos & Csizer 2008; Kozaki & Ross 2011; Lamb 2009; Ryan 2009; Taguchi, Magid & Papi 2009; and Ushioda 2009). The L2 self-system developed into

a critical factor in explaining individual learner interest or motivation, resulting success or lack thereof in L2 learning (Campbell & Storch 2011). This previous research is essential to the present study, as we seek to better understand L2 learners' feelings and motivations in many of these areas.

As new issues in teaching and learning foreign languages emerge, we face another renewal of interest in the research of L2 learners' motivation. In the 2010s, university L2 learners are confronted with high educational costs, a shifting economy and a competitive job market. The ways in which students meet these challenges, what influences their motivation (job search, marketability, employability, cultural ties to the language) to study Spanish and how motivation to study Spanish may vary according to the geographic region of the student are areas that this study investigates. Finally, we discuss how current learner motivations may affect university L2 course offerings, curriculum and pedagogy. Therefore, it is time to revisit a variety of topics surrounding L2 learner motivation and update previous research to today's reality in foreign language education. In this study, we will update previous research and compare how, if at all or to what degree, students' motivation to take Spanish has changed.

Two previous studies (McEwan & Minkle 1979; Ely 1986) about L2 motivation emerged to us as particularly interesting and relevant in terms of research questions, data collection instruments used and presentation of results. While reviewing the literature, we found ourselves wondering how current university students of Spanish in the U.S. would answer these researchers' questions about motivation and how these answers may differ from those of 1979 and 1986, respectively. We believe the information presented in these specific articles should be reexamined in the 2010s, particularly in light of the rise in the Hispanic population in the U.S. and subsequent increase in exposure to the Spanish language and its cultures. As a consequence, it is essential to update the possible changing attitudes of U.S. university students toward their motivations to learn Spanish. Therefore, the present study developed into a conceptual replication study based on the works of McEwan & Minkle (1979) and Ely (1986).

### 3. Methods

#### 3.1 Instruments

For this study, we employed a modified version of the questionnaire used in McEwan & Minkle (1979) and Ely (1986). Before distributing our data collection instrument for the present study, we conducted an informal pilot study with our own respective language students in order to discover possible issues with the questionnaire and to modify it as necessary. Part I included a survey with a

5-point Likert scale, with the alternatives labeled: “strongly disagree,” “disagree,” “neither agree nor disagree,” “agree,” and “strongly agree.” Participants were asked to indicate their level of agreement with varying types of motivation categories. To be consistent with the conceptual replication of McEwan & Minkle (1979) and Ely (1986), we were careful to include instrumental motivations, such as a desire to get a job, in addition to integrative ones, like connecting with the L2 culture. Part II of the survey asked students to rank their top five reasons (out of nine) for taking Spanish in order to determine the strength of motivation.

This two-part instrument provided results that could be compared and contrasted with those from previous research. However, we were left wondering if participants had additional motivations or information that they could share that did not emerge from the structure of the questionnaire. Ely (1986: 32) suggested that

Further research is necessary to assess the relationship between type of motivation and strength of motivation among other groups of learners. The first step in such research could be the sort of initial open-ended survey (of type of motivation) employed in the present research. Such surveys are likely to lead to the discovery of motivational patterns that theory alone might not reveal.

Therefore, due to the relatively small sample size of the present study and the notion that few researchers have heeded Ely’s suggestion, we took advantage of the opportunity to obtain qualitative data through open-ended questions. Hence, in the data collection instrument used in the present study, Part III included open-ended questions in search of potential additional motivators for students that were possibly left uncovered in the questionnaire. Also, due to the sample size, this study was designed to serve as a pilot study for a future project with more participants.

Thus, in the present study, the main instrument was a survey presented to students of Spanish who were enrolled in elementary-level Spanish courses at two universities. After providing biographical information (age, name of the class in which enrolled, major, minor and year in college), students filled out three sections of the survey to determine their motivation for studying the Spanish language. The students did not provide their names and were told that their participation was voluntary.

### 3.2 University demographics

The university in the American Midwest has approximately 3,000 undergraduate students and about 1,700 graduate students. 64% percent female and 36% male. 89% of students who attend this university hail from the state in which the university is located. The race/ethnicity breakdown of the students from this university is as follows: 66% white; 9% African American; and 17% Latino/Hispanic. The university in the American South has about 6,500 undergraduate students and

about 650 graduate students. 63% of students are female and 37% are male. 92% of students are from the state where the university is located. The race/ethnicity breakdown of the students from this university is as follows: 68% white; 20% African American, 3% Latino/Hispanic. It should be noted that the smaller, private university in the Midwest has a larger Hispanic student population and access to Spanish speakers in the community (see Fafulas et al., this volume, for discussion of the influence of speakers' contact with the speech community). This is partially due to its proximity to Chicago, which now has the fifth largest Hispanic metropolitan population in the U.S. at 1.9 million people (Motel & Patten 2012). Furthermore, the surrounding communities (within 40 miles of Chicago city limits) of this university are now 40% Hispanic, with 63% of new residents in the last ten years being of Hispanic descent (Tareen 2011).

It should also be mentioned that the larger university is public and located in a rural area of the American South. This area of the country has also seen an increase in the number of Hispanic residents, though not as great as that of a large city located near the Midwestern university. Census data reveals that between 2000 and 2010 the Hispanic population of the state where the southern university is located grew 79% and that more than half of this population lives in the southeastern part of the state (Hispanic Leadership Network, 2012), where data for the current study were collected. Therefore, these universities were selected in part because they are both in states and geographical areas that have seen significant increases in the Hispanic population. Furthermore, we chose these universities as we were given full access to the students as study participants. Also of interest, the private Midwestern university, while smaller in enrollment, offers a Spanish major and minor, while the larger southern university offers only a Spanish minor.

### 3.3 Participants

The participants for this study were college students enrolled in undergraduate Spanish courses at four-year universities in the U.S. Thirty-four students attended the private university in a suburb of a large city in the American Midwest, and thirty-five were students at a public university in a small city in the American South. Overall, there were 54 students whose survey results were included in this study. The students were enrolled in elementary-level Spanish courses. These beginner courses are university or general education requirements at the two institutions in question and enroll students from diverse academic majors and backgrounds.

The average age of the students was 20 years of age. Our sample was overwhelmingly female (approximately 63%) and 31% were male. 6% of participants did not disclose their age. Roughly 41% of participants were first-year students, 24% were second-year, 20% were third-year and 13% were fourth-year (2% did



not specify their years in college). The vast majority of students were not Spanish minors (78%), while only 22% said they had declared a Spanish minor. Fifteen students' responses were unable to be calculated for the purposes of this study as their surveys were incomplete or filled out incorrectly. We will address this in the Limitations section of this study.

3.4 Procedures

Students completed the anonymous three-part survey described above about their foreign language study and motivations to study Spanish at the university level. We analyzed the results from these surveys, classified them by motivation type and compared them according to several variables. The results of these surveys are found in the following section.

4. Results

In this section, the results from our surveys will be presented in chart form and discussed briefly. To follow the conceptual replication study in the form of McEwan & Minkle (1979) and Ely (1986), we classify the motivations as integrative and instrumental. As previously mentioned, certain motivations, such as “the desire to communicate with Spanish speakers,” could be labeled both integrative and instrumental depending on the context. Tables 1 and 2 describe the motivations of the students for studying Spanish, while Tables 3 and 4 compare the students' responses by their location (Midwestern U.S and Southern U.S). The significance and implications of the results will be discussed in §5.

4.1 Motivation to study Spanish (part I)

First, students were asked to describe to what extent the various types of motivation influenced them in their choice to study Spanish. The results are in Table 1.

Table 1. Motivations for taking Spanish (*N* = 54)

Question	Mean
May be helpful in my future career	4.52
Mentally challenging and provides mental exercise.	4.37
Spanish is an important language in the world.	4.35
May make me a more qualified job candidate.	4.31
Able to speak more languages than just English.	4.28

Table 1. (*Continued*)

Question	Mean
Foreign language study is part of a well-rounded education.	4.17
Able to converse with Spanish speakers in the U.S.	4.13
Use Spanish when I travel to a Spanish-speaking country.	4.00
Learn about another culture to understand the world better.	3.69
Interested in Hispanic culture, history, music, film, or literature.	3.52
Able to use it with Spanish-speaking friends.	3.43
Fulfills a general education course.	3.24
Studying a foreign language is a requirement for my major.	3.19
Needed for study abroad.	2.83
Helps in understanding English grammar better.	2.57
Spanish is less challenging than other foreign languages.	2.57
Interest in my own Hispanic heritage.	2.50
Communicate with relatives in Spanish.	2.50

The answer most commonly selected (4.52 out of 5) as being highly motivating was “May be helpful in my future career” (instrumental motivation). “May make a more qualified job candidate” (instrumental) was also a common response (4.31). Interestingly, “Spanish is an important language in the world” (instrumental) and “Able to converse with Spanish speakers in the U.S” (instrumental) ranked high (4.35 and 4.13, respectively). However, the desire to “Communicate with relatives in Spanish” (integrative) ranked last (2.50) along with “Interest in my own Hispanic heritage” (integrative). “Mentally challenging and provides mental exercise” was high at 4.37 and was a strong integrative motivation, which was not originally expected to be so high. Due to the fluctuating job market, it was interesting that integrative motivators, such as a desire to have a well-rounded education (4.17%) had such a strong presence.

The open-ended individual student comments from the third section of our survey added to the depth of this study’s results. Some students did mention their families as (integrative) motivators for studying Spanish: “my heritage is Hispanic, so I might as well. My family heritage is Spanish.” One student wrote, “I’ve always wanted to learn a second language, and I think the culture and language is interesting.” Several students also cited an increase in the number of Hispanics in the U.S. as a reason to study Spanish: “more and more people who migrate to the U.S will speak Spanish as their primary language” as well as “because there are more Spanish speaking people here than any other language.”

One student mentioned being motivated by the popularity of Spanish world-wide: “because a lot of people speak it around the world.” As the results of the first section of the questionnaire revealed, a desire to find gainful employment (instrumental motivation) was in the forefront of students’ minds, as indicated in these comments: “It would have made getting a job in the field of nursing easier” and “I plan on becoming a teacher and it will help me to converse with Spanish families easier.”

4.2 Motivation to study Spanish (part II)

Students were then asked to rank their top reasons and motivations for studying Spanish (Table 2).

Table 2. Top reasons for taking Spanish (N = 54)

Rank	Reason
1	More marketable on the job search or will find a better job.
2	A personal goal to learn another language.
3	Spanish will be an important language in the U.S.
4	Travel or study abroad in a Spanish-speaking country.
5	A requirement for major or for general-education hours.
6	Make more money if bilingual.
7	Learn more about Spanish-speaking countries.
8	Spanish is my minor or major.
9	Family speaks Spanish.

Again, a desire to be more marketable (instrumental factor) on the job market proved to be their main motivation. Of interest is that a combined integrative-instrumental goal, a desire to be bilingual, was the second motivation, followed by “Spanish will be an important language in the U.S.” (instrumental) and a desire to speak Spanish while traveling (instrumental). Learning Spanish to speak with family was ranked lowest by participants, while learning Spanish because it was the student’s minor or major was the second lowest.

4.3 Motivation to study Spanish – geography (part I)

We also examined the connection between geographic area of the U.S. and the motivation to study Spanish. When comparing the two groups of students by geographic region, the results were also quite telling (Table 3).

Table 3. Motivations for taking Spanish by location

Southern U.S. students (N = 40)		Midwestern students (N = 14)	
	Mean	Question	Mean
Mentally challenging and provides mental exercise.	4.73	May be helpful in my future career	4.57
May be helpful in my future career	4.50	Able to speak more languages than just English.	4.43
May make a more qualified job candidate.	4.35	Spanish is an important language in the world.	4.43
Spanish is an important language in the world.	4.33	Able to converse with Spanish speakers in the U.S.	4.29
Able to speak more languages than just English.	4.23	May make a more qualified job candidate.	4.21
Foreign language study is part of a well-rounded education.	4.20	Learn about another culture to understand the world better.	4.14
Able to converse with Spanish speakers in the U.S.	4.08	Use Spanish when I travel to a Spanish-speaking country.	4.07
Use Spanish when I travel to a Spanish-speaking country.	3.98	Foreign language study is part of a well-rounded education.	4.07
Interested in Hispanic culture, history, music, film, or	3.55	Fulfills a general education course.	3.50
literature. Learn about another culture to understand the world	3.53	Interested in Hispanic culture, history, music, film, or literature.	3.43
better. Able to use it with Spanish-speaking friends.	3.45	Needed for study abroad	3.43
Studying a foreign language is a requirement for my major.	3.35	Mentally challenging and provides mental exercise.	3.36
Fulfills a general education course.	3.15	Able to use it with Spanish-speaking friends.	3.36
Needed for study abroad	2.63	Helps in understanding English grammar better.	2.86
Spanish is less challenging than other foreign languages.	2.58	Communicate with relatives in Spanish.	2.79
Interest in my own Hispanic heritage.	2.53	Studying a foreign language is a requirement for my major.	2.71
Helps in understanding English grammar better. Communicate	2.48	Spanish is less challenging than other foreign languages.	2.57
with relatives in Spanish.	2.40	Interest in my own Hispanic heritage.	2.43

“Mentally challenging and provides mental exercise” (integrative) was the main motivator for the Southern U.S. students (4.725 out of 5), yet it was far less important to the Midwestern U.S. students (3.36). Similarly, open-ended student comments also mentioned less common motivators such as “I had a free elective and thought Spanish would be fun and interesting - which it was!” (integrative; Midwest). Citing a desire to connect with other cultures, one student wrote, “I think it is important to learn another language to break the cultural barrier among different cultures” (integrative; South). An interest in finding gainful employment (instrumental) and earning a better salary (instrumental) were important to both sets of students. The Midwestern students agreed with the southern students that “Interest in my own Hispanic Heritage” (integrative) was a less important motivation (2.4 and 2.525, respectively). “Communicate with relatives in Spanish” (instrumental), however, was a slightly higher motivation to the Midwestern students than to the Southern students (2.79 and 2.4, respectively). The emphasis placed on learning Spanish for monetary gain (instrumental) was also apparent in the comments of students: “it would have made getting a job in the field of nursing easier” (Midwest) and “I plan on becoming a teacher and it will help me to converse with Spanish families easier” (South). Another student added, “employers will see an achievement in Spanish and possible [sic] consider you over someone without Spanish” (South). Finally, a student mentioned both instrumental (job) and integrative (culture) motivators:

The job market for plain history majors is not very promising. Learning Spanish and going for Latin History narrows the market but is a less common field. It will be easier to get a job. Also, learning Spanish opens one up to a continent rich in culture and interesting historical facts (South).

Taking Spanish classes as a motivation for studying abroad was not a primary motivator for either group. This may be due to neither university having a large percentage of students enrolled in study abroad programs. Neither school has a large variety of study abroad offerings, so perhaps this motivation would vary at another university.

#### 4.4 Motivation to study Spanish – geography (part II)

Next, students were asked to rank their motivations to study Spanish. The results from the surveys completed by students at the university in the South were then compared to those of students at the university in the Midwest. These results are found in Table 4.

**Table 4.** Top reasons for taking Spanish by location

Southeastern U.S. ( <i>N</i> = 40)		Midwestern U.S. ( <i>N</i> = 14)	
Rank	Reason	Rank	Reason
1	More marketable on the job search or will find a better job.	1	More marketable on the job search or will find a better job.
2	A personal goal to learn another language.	2	Travel or study abroad in a Spanish-speaking country.
3	Spanish will be an important language in the U.S.	3	Make more money if bilingual.
4	A requirement for major or for general-education hours.	4	Spanish will be an important language in the U.S.
5	Travel or study abroad in a Spanish-speaking country.	5	A personal goal to learn another language.
6	Make more money if bilingual.	6	Learn more about Spanish-speaking countries.
7	Learn more about Spanish-speaking countries.	7	Spanish is my minor or major.
8	Spanish is my minor or major.	8	A requirement for major or for general-education hours.
9	Family speaks Spanish.	9	Family speaks Spanish.

When ranking their motivations, the highest and lowest motivations were the same, despite the varying demographics of the two geographic regions where the data was collected. Both groups cite “Spanish will be an important language in the U.S” (instrumental) highly, while a desire to “Learn more about Spanish-speaking countries” (integrative) was low for both groups.

## 5. Discussion

Our results offer insight into the various types of motivation of university students of Spanish. While some of these results were in agreement with previous findings, others contrast with earlier results. Many, particularly the interest of students in finding quality employment, have pedagogical implications. Other results, particularly the lack of students studying Spanish because of a family connection, were also revealing and could have implications in the classroom. In this section, we discuss how these results relate to previous research and their pedagogical implications. Finally, we address the limitations of this study that scholars might consider for future study.

## 5.1 Research implications

Our results have updated previous research in several ways, and have added to the discussion of current student motivation to learn Spanish. It is clear from all previous research reviewed and from the results of the present study that learners of Spanish are motivated by both integrative and instrumental factors. Since the instruments used in McEwan & Minkle (1979) and Ely (1986) were essential in informing our questionnaire and instrument in this conceptual replication study, we can discuss in particular how our participants' motivation in 2014 compared with or differed from their participants' motivation in the 1970s and 1980s, respectively. In McEwan & Minkle (1979: 92), the most common reasons (indicated by at least 65% of respondents) for college students learning Spanish included "it would be fun to learn and know a foreign language" (integrative/72%), "to visit a foreign country" (instrumental/73%) and "help me understand other people better" (integrative/68%). In the present study, students in both the Midwest and the South were also motivated by integrative and instrumental factors for learning Spanish (see Table 1). However, they were much more likely to choose factors concerning employment or career than students in McEwan & Minkle's study (1979: 92): "May be helpful in my future career"; "Mentally challenging and provides mental exercises"; "Spanish is an important language in the world"; and "May make a more qualified job candidate." In the present study, top reasons (strength of motivation; see Table 2) again showed integrative and instrumental factors, but students were more focused on motivation surrounding their careers: "More marketable on the job search or will find a better job" and "A personal goal to learn another language". Our results were similar to those of Ely (1986), as both integrative and instrumental motivators were revealed as important factors in learning Spanish.

With regard to the influence of geography, there was not much difference between the students at the two universities and that economically influenced motivators were most important to both groups. The Midwestern students only slightly favored learning Spanish to "Communicate with relatives in Spanish" over the Southern students, despite the increasing number of Hispanics in the Chicago area. As discussed in our results section, adding a section to our instrument with open-ended student responses did contribute to our findings (as suggested by Ely 1986). Students mentioned more often than not a motivation to learn Spanish because of an employment/career-oriented (instrumental) reason or because Spanish is an important language in the world and because they wanted to have the ability to communicate in the language (cf. Lynch, this volume, for a discussion of possible changing motivations for English language acquisition in Spanish-speaking areas). Many students mentioned that they were in contact with Spanish speakers and the Spanish language in the U.S. and that taking Spanish classes was

good preparation and training for future employment. The motivational factors of these students bring to the forefront of our discussion many pedagogical implications for the future of language teaching and learning.

## 5.2 Pedagogical implications

Our results have a number of pedagogical implications. Since the results revealed that students are interested in earning money and finding employment, Spanish courses that emphasize special interests or topics might be popular, beneficial and fulfilling for students. These special-topics courses, such as Spanish for Educators, Business Spanish, Medical Spanish, or Spanish for Social Work, could also lead to interdisciplinary studies and work among various departments. Domestic internships, volunteer or service-learning classes and other experiential learning opportunities using Spanish are also important considerations in current Spanish curriculum discussions. These experiences can connect students with Spanish speakers in the community and provide necessary contact with Hispanic cultures, either of which could fulfill integrative and/or instrumental motivational orientations. Instructors and administrators might also mention possible monetary gains and the increased ease in finding employment or marketability that knowing Spanish can provide.

A desire to learn Spanish for travel or study abroad, while not a primary factor, was important to a number of participants. In addition to offering courses that prepare students for study abroad, these study abroad programs might also offer opportunities of experiential learning abroad (i.e., internships, service learning, volunteer projects) so that two motivators (a desire to travel and find employment) are linked.

## 5.3 Limitations and future study

One limitation was the relatively small size of our participant group (54 students); a larger sample would certainly be more informative. Also, to better understand the effect of the geographic location of the universities that we sampled, not only would adding more areas be beneficial, but a more balanced sample from each area would also be ideal.

Students enrolled in upper-level Spanish courses were not polled since they represented a small segment of the overall enrollment in Spanish courses. If the study were to be conducted again, it would be important to poll both lower and upper-level Spanish students to gauge possible differences in motivational factors due to course level.

Finally, several surveys were completed incorrectly and could not be included in the study results. As this is a replication study that seeks to use instruments



comparable to those used by Ely (1986) and McEwan & Minkle (1979), we did not completely alter the instruments. In future studies, it is advisable that the directions of the survey be repeated out loud. It is also essential to verify that students have completed the questionnaire correctly each time a participant turns one in, as we did have a few surveys that were not completed correctly.

There are many ways in which the present study might be adapted for future investigations. First, as mentioned above, a larger sample size including students from various parts of the country who are studying different foreign languages would provide valuable information. Also, specifically knowing which students were heritage speakers would have been beneficial since several students of Latino/Hispanic descent disclosed that the desire to hear a variety of Spanish other than their own motivated them to enroll in a university Spanish course.

Studying abroad was a low-ranking motivator overall in this study. While the universities sampled in this project have a limited number of study abroad programs, it was expected that a desire to study abroad or travel would be more of a motivating factor. The response could be due to the rising cost of travel or economic hardships; however, in future study, we would allow for more room for students to comment, or we would ask a specific open-ended question about study abroad.

Furthermore, it would be interesting to investigate issues that have emerged in previous research: Did students' attitudes toward the target language or culture affect their motivation? Do students' attitudes toward their instructor change their motivation to learn the language? An achievement question in Spanish (i.e., students' grades or success in the class or language) would have been interesting for the investigation into how achievement affects motivation and vice versa (cf. Hernandez 2008; Campbell & Storch 2011). Dörnyei (2009), along with his other studies, suggests that integrative motivation orientations become more internalized motives and are more likely to have a positive impact on L2 learning. Recent studies developing from Dörnyei and those that have explored the L2 self and learner self-efficacy (cf. Grenfell & Harris 2013; Iaccarino 2012; Kormos & Csizer 2008; Kozaki & Ross 2011; Lamb 2009; Ryan 2009; Taguchi et al. 2009; Ushioda 2009) could be expanded upon in future research by adding questions to the present study that examine the achievement, or learner-perceived success or lack thereof, in L2 learning.

Other areas of previous research that could be elaborated in the future include specific study of heritage learners' motivation to study Spanish (Hsieh 2008), gender differences in motivation to study Spanish (Kissau et al. 2010), technology's effect on student motivation (Negueruela-Azarola 2011; Lynch, this volume) and changes in motivation and attitude throughout a course (Gardner et al. 2004). Finally, further research is needed to identify suggestions and motivational

strategies for instructors (Dörnyei 2001; Mercè Bernaus & Gardner 2009), which could be of use for both departments and professional development activities for faculty.

## 6. Conclusions

Our results correlate with a shift in the American economy that has led students to look to university courses as a marketability tool to possibly improve their chances of obtaining employment after graduation. If educators are able to better understand what draws students to enroll in their classes, their response can have two outcomes. First, educators might incorporate special topics into curricula to attract students based on their motivations and interests. Also, instructors may improve the retention of students in their Spanish courses if they are aware of the students' motivations. As many institutions of higher learning face budget cuts and staff reductions, attracting and retaining students is of utmost importance. It is our hope that the results of this study will lead to increasing numbers of college students enrolling and benefitting from the academic and cultural awareness that studying a foreign language brings.

As student populations and demographics continue to change and contact with the Spanish language and Spanish-speaking cultures in the U.S. rises, educators must constantly be aware of what motivates student learning and goal orientation. An educator who is aware of the current motivational factors as to why students have chosen a course, a minor/major or a career path can better serve the students and academic programs. Understanding students' integrative and instrumental motivations also can aid in innovative course and program development, such as experiential learning opportunities, that will attract and retain students to particular courses or subject areas.

## References

- Campbell, Elizabeth & Neomy Storch. 2011. "The Changing Face of Motivation: A Study of Second Language Learners' Motivation over Time". *Australian Review of Applied Linguistics* 34:3.166–192.
- Cesar, Stephen. 2011, March 24. "Hispanic Population Tops 50 Million in the U.S". *Los Angeles Times*. Retrieved from <http://articles.latimes.com/2011/mar/24/nation/la-na-census-hispanic-20110325>.
- Deci, Edward L. 1992. "The Relation of Interest in the Motivation of Behavior: A Self-Determination Theory Prospective". *The Role of Interest in Learning and Development* ed. by K. Ann Renninger, Suzanne Hidi & Andreas Krapp, 43–70. Hillsdale, N.J.: Lawrence Erlbaum.

- Deci, Edward L., Robert J. Vallerand, Luc G. Pelletier & Richard M. Ryan. 1991. "Motivation and Education: The Self-Determination Perspective". *Educational Psychologist* 26.325–346. doi:10.1080/00461520.1991.9653137
- Dörnyei, Zoltán. 1994. "Motivation and Motivating in the Foreign Language Classroom". *Modern Language Journal* 78.273–284. doi:10.1111/j.1540-4781.1994.tb02042.x
- Dörnyei, Zoltán. 1998. "Motivation in Second and Foreign Language Learning". *Language Teaching* 31.117–135. doi:10.1017/S026144480001315X
- Dörnyei, Zoltán. 2001. *Teaching and Researching Motivation*. Essex: Longman.
- Dörnyei, Zoltán. 2009. "The L2 Motivational Self-System". *Motivation, Language Identity and the L2 Self* ed. By Zoltán Dörnyei & Ema Ushioda, 9–42. Bristol: Multilingual Matters.
- Ely, Christopher M. 1986. "Language Learning Motivation: A Descriptive and Causal Analysis". *The Modern Language Journal* 70.28–35. doi:10.1111/j.1540-4781.1986.tb05240.x
- Gardner, Robert C. 1985. *Social Psychology and Second Language Learning: The Role of Attitudes and Motivation*. Baltimore: Edward Arnold.
- Gardner, Robert C. & Richard Clément. 1990. "Social Psychological Perspectives on Second Language Acquisition". *Handbook of Language and Social Psychology* ed. by Howard Giles & William Peter Robinson, 495–517. Chichester, England: Wiley.
- Gardner, Robert C. & Peter D. MacIntyre. 1993. "A Student's Contributions to Second-language Learning. Part II: Affective Variables". *Language Teaching* 26:1.1–11. doi:10.1017/S0261444800000045
- Gardner, Robert C., Richard Clément, Padric C. Smythe & C.L. Smythe. 1979. *The Attitude/Motivation Test Battery-Revised Manual*. London, Canada: University of Western Ontario.
- Gardner, Robert C. & Wallace E. Lambert. 1972. *Attitudes and Motivation in Second Language Acquisition*. Rowley, Mass.: Newbury House.
- Gardner, Robert C., Anne-Marie Masgoret, Jeff Tennant & L. Mihic. 2004. "Integrative Motivation: Changes During a Year-Long Intermediate-Level Language Course". *Language Learning* 54.1–34. doi:10.1111/j.1467-9922.2004.00247.x
- Grenfell, Michael & Vee Harris. 2013. "Making a Difference in Language Learning: The Role of Sociocultural Factors and of Learner Strategy Instruction". *The Curriculum Journal* 24:1.121–152. doi:10.1080/09585176.2012.744326
- Hernandez, Todd A. 2008. "Integrative Motivation as a Predictor of Achievement in the Foreign Language Classroom". *Applied Language Learning* 18.1–15.
- Hispanic Leadership Network. 2012, March 23. "Top 10 Facts about Hispanics in Louisiana". Retrieved from <http://hispanicleadershipnetwork.org/top-10-facts-about-hispanics-in-louisiana/#sthash.5d9cGRhm.dpuf>.
- Hsieh, Pei-Hsuan. 2008. "Why are College Foreign Language Students' Self-Efficacy, Attitude, and Motivation so Different?" *International Education* 38.76–94.
- Iaccarino, Antonio R. 2012. "The Role of Motivation and Controversial Conceptual Material in Foreign Language Classrooms". *Journal of Language Teaching and Research* 3:3.430–438. doi:10.4304/jltr.3.3.430-438
- Kissau, Scott P., Lan Quach Kolano & Chuang Wang. 2010. "Perceptions of Gender Differences in High School Students' Motivation to Learn Spanish". *Foreign Language Annals* 43.703–721. doi:10.1111/j.1944-9720.2010.01110.x
- Kormos, Judit & Kata Csizer. 2008. "Age Related Differences in the Motivation of Learning English as a Foreign Language: Attitudes, Selves and Motivated Learning Behavior". *Language Learning* 58.327–355. doi:10.1111/j.1467-9922.2008.00443.x

- Kozaki, Yoko & Steven Ross. 2011. "Contextual Dynamics in Foreign Language Learning Motivation". *Language Learning* 61:4.1328–1354. doi:10.1111/j.1467-9922.2011.00638.x
- Lamb, Martin. 2009. "Situating the L2 Self: Two Indonesian School Learners of English". *Motivation, Language Identity and the L2 Self* ed. by Zoltán Dörnyei & Ema Ushioda, 229–247. Bristol: Multilingual Matters.
- McEwan, Angela & Roger Minkle. 1979. "Why Am I Learning Spanish? An Evaluation of Student Response". *Hispania* 61.89–97. doi:10.2307/340083
- Mercè Bernaus, Annie-Wilson & Robert C. Gardner. 2009. "Teachers' Motivation, Classroom Strategy Use, Students' Motivation and Second Language Achievement". *Porta Linguarum* 12.25–36.
- Motel, Seth & Eileen Patten. 2012, September 19. "Characteristics of the 60 Largest Metropolitan Areas by Hispanic Population". *Pew Research Center: Hispanic Trends*. Retrieved from <http://www.pewhispanic.org/2012/09/19/characteristics-of-the-60-largest-metropolitan-areas-by-hispanic-population/>.
- Negueruela-Azarola, Eduardo. 2011. "Changing Reasons as Reasoning Changes: A Narrative Interview on Second Language Classroom Motivation, Telecollaboration, and the Learning of Languages". *Language Awareness* 20.183–201. doi:10.1080/09658416.2011.570348
- Paris, Scott G. & Julianne C. Turner. 1994. "Situating Motivation". *Student Motivation, Cognition, and Learning* ed. by Paul R. Pintrich, Donald R. Brown & Claire Ellen Weinstein, 213–237. Hillsdale, N.J.: Lawrence Erlbaum Associates.
- Ryan, Stephen. 2009. "Self and Identity in L2 Motivation in Japan: The Ideal L2 Self and Japanese Learners of English". *Motivation, Language Identity and the L2 Self* ed. by Zoltán Dörnyei & Ema Ushioda, 98–119. Bristol: Multilingual Matters.
- Taguchi, Tatsuya, Michael, Magid & Mostafa Papi. 2009. "The L2 Motivational Self System among Japanese, Chinese and Iranian Learners of English: A Comparative Study". *Motivation, Language Identity and the L2 Self* ed. by Zoltán Dörnyei & Ema Ushioda, 66–97. Bristol: Multilingual Matters.
- Tareen, Sophia. 2011, June 8. "Latino Population in Chicago Suburbs Explodes, Changing Makeup of Region". *The Huffington Post*. Retrieved from [http://www.huffingtonpost.com/2011/04/08/latino-population-in-chic\\_n\\_846894.html](http://www.huffingtonpost.com/2011/04/08/latino-population-in-chic_n_846894.html).
- Ushioda, Ema. 2009. "A Person-in-context Relational View of Emergent Motivation, Self and Identity". *Motivation, Language Identity and the L2 Self* ed. by Zoltán Dörnyei & Ema Ushioda, 215–228. Bristol: Multilingual Matters.
- Vallerand, Robert J. 1997. "Toward a Hierarchical Model of Intrinsic and Extrinsic Motivation". *Advances in Experimental Psychology* 29.271–360. doi:10.1016/S0065-2601(08)60019-2



## *Espero estén todos*

# The distribution of the null subordinating complementizer in two varieties of Spanish

Adrián Rodríguez Riccelli

University of Texas at Austin

This study adopts a quantitative methodology to investigate variation between null and overt expressions of Spanish complementizer *que*. I focus on two varieties of Spanish from Mexico City and Los Angeles, California. This research uses a corpus linguistics-inspired methodology, relying on Twitter as a database to extract tokens and Rrbul (Johnson 2008) to perform the multivariate analysis. I extracted and analyzed 1,505 tokens and coded for 9 linguistic and 2 extra-linguistic factors. I showed the internal factors ‘verb modality’ and ‘embedded subject status’ to significantly affect the alternation, favoring the null option. Contrary to what was expected, the external factor ‘city’ did not have any significant effect on the distribution. I discuss the results in light of formal and sociolinguistic studies on the nature of subordinated constructions in Spanish (Brovetto 2002; Etxepare 1996; Subirats-Rüggeberg 1987), English (Tagliamonte & Smith 2005) and Spanish in contact with English (Silva-Corvalán 1993, 1994, 1998).

## 1. Introduction

Unlike languages that allow for high levels of optionality in the overt expression of certain complementizers, the Spanish subordinating complementizer *que* is generally thought to be limited to only overt realizations. However, it has been observed that in highly polite, formal, and written contexts, *que* may be omitted, such as with the verb *rogar* in (1).

- (1) *Estoy enfermo, le ruego Ø no fume*  
 COP.1SG sick ACC.2SG plead.1SG COMP NEG smoke.2SG.SBJV  
 “I am sick, I plead (that) you not smoke.”

(“¿Es peligroso fumar en un Taxi?”, ABC 1977: 49)

This phenomenon has long been debated in the syntactic-theoretical literature on Spanish, where *que*-drop has often been investigated as a means of exploring the architecture of the CP or left periphery (Antonelli 2012; Broveto 2002; Etxepare 1998; Torrego 1983; Villa-García 2014). Alongside these formal approaches to the study of *que*-drop, sociolinguistic analyses of the phenomenon have been carried out by Silva-Corvalán (1994, 1998, 2007), who attested the presence of null complementizers in the Spanish of Los Angeles. In these studies, the author suggests that *que*-drop appears to be driven by contact with English. However, she clarifies that the reason for its presence in Los Angeles Spanish would not be due to a direct borrowing or calquing of English syntax but rather to a relaxation of Spanish pragmatic constraints. She concludes that the null complementizer, formerly encountered only in certain formal contexts in monolingual Spanish, as in (1) for example, is now spreading to new pragmatic contexts in the speech of bilinguals. She further argues that the emergence of null complementizers in Los Angeles Spanish is characteristic of a larger language-internal process of change that is catalyzed by the dynamics of language contact where the contact language, in this case English, demonstrates parallel structures in its own grammar.

Mode and style are also implicated in the variability of the expression of subordinating *that* in English. High rates of *that*-drop are found in spoken and informal registers of English, while the overt option is preferred in written modes and in more formal registers (Elseness 1984). In oral modes, the highest rates of null complementizers are found with matrix collocations such as *I think*, *I guess*, and *you know* (Tagliamonte & Smith 2005, with references). The variation is conditioned by a number of syntactic factors, including the status of the matrix and embedded subjects, the presence of material intervening between the two clauses and co-referentiality between the two subjects, among other factors (Kearns 2007; Tagliamonte & Smith 2005).

The present study builds on previous work to cast light on the nature of *que*-drop in two dialects of Spanish, those of Mexico City and Los Angeles, California. I place particular emphasis on instances of *que*-drop in informal registers of Spanish, since we have already seen evidence in historical records and print media (Subirats-Rüggeberg 1987) that *que*-drop has long been an option in Spanish subordinating structures. To this end, data were drawn from Twitter, a social media platform whose language tends to be informal in nature and whose language output lies somewhere between oral and written modes of language. In addition, I seek to reveal a detailed account of the syntactic factors in the clauses adjacent to the subordinating complementizer, to identify the contexts in which *que*-drop most often emerges. In comparing the Mexico City and Los Angeles dialect regions – two distinct but related varieties of Spanish – I seek to determine if Los Angeles Spanish has undergone syntactic transfer, borrowing or calquing or experienced other ‘indirect’ contact effects from Southern California English. In turn, this line

of research contributes to a broader questions in contact linguistics, the question of whether or not syntactic transfer in situations of language contact is possible at all. Furthermore, it allows one to explore further the structural properties of the Spanish complementizer system using an empirical approach. This investigation adopts a large-scale quantitative methodology that applies corpus-research-inspired methods in line with the work carried out on English *that*. In particular, it relies on the use of Twitter as a database to extract tokens and Rbrul (Johnson 2008), a statistical software package, to perform the quantitative analyses.

This paper consists of six sections. Section 1 summarizes the research on *que*-drop that has been carried out on Spanish and Spanish in contact with English, detailing the numerous syntactic factors proposed for where *que*-drop should be possible. Section 2 provides an overview of the main findings on variable complementizer drop in English, revealing numerous syntactic and extra-linguistic variables that coalesce to allow for very high rates of *that*-drop in spoken English, and a dramatic increase in the rates of *that*-drop historically. Section 3 presents the research questions posed in the current study. Section 4 describes the methodology adopted, in particular the use of software to extract data from the social media platform Twitter within a geographically delimited region. Section 5 presents results from the statistical analyses showing two syntactic factors, verb modality and embedded subject position, to be the strongest predictors of *que*-drop. Section 6 discusses the broader implications showing that in contexts of *well-wishing* and expressing *hopes and desire for a beneficial outcome*, matrix verb *esperar* and its accompanying bi-clausal complex is undergoing a process of grammaticalization whereby the subordinating structure is eroded causing *que* to be dropped, a property that serves to mark *volitional* or *boulomaic modality*.

## 2. Null complementizers in Spanish

Most investigations of *que*-drop before finite subordinate clauses in Spanish converge on the finding that the option is limited to only a small set of lexical verbs. While this class of verbs has been given a variety of labels depending on the author, the conceptual relationship shared between the verbs seems to justify their grouping into a single set based on two or three forms of modality.<sup>1</sup> Loosely, this class includes verbs such as: *suponer* ‘to suppose’, *dudar* ‘to doubt’, *parecer*

---

1. Cattell (1977) as well as Torrego (1983) and Exteparre (1998) label these ‘stance predicates’ while Broveto (2002) calls them ‘verbs of propositional attitude’, ‘volitional verbs’, verbs of desire’, and ‘lamentar-like verbs’.



“to seem”, *lamentar* “to lament”, *preocuparse* “to worry”, *alegrarse* “to be glad”, *querer* “to want”, *desear* “to desire”, and *esperar* “to hope”. Examples from Broveto (2002: 34) follow in (2)–(3).

- (2) *Lamento* Ø *no estás contenta*  
 lament.1SG COMP NEG are.2SG.SBJV happy  
 “I am sorry that you are not happy.”
- (3) *Espero* Ø *se solucionen pronto los problemas*  
 hope.1SG COMP 2P solve.2P.SBJV soon DET problems  
 “I hope that your problems are solved soon.”

On the other hand, some predicates have been claimed to disallow *que*-drop, such as verbs of saying *decir* “to say”, *repetir* “to repeat”; factive verbs *confesar* “to confess”, *jurar* “to swear”, and *admitir* “to admit”, as illustrated in (4)–(5) (Broveto 2002: 34).

- (4) \**Dijo* Ø *llegó tarde a la reunión*  
 say.1SG COMP arrive.3SG late to DET reunion  
 “He said that he arrived late to the reunion.”
- (5) \**Confieso* Ø *he mentido repetidamente*  
 confess.1SG COMP have.1SG lied repeatedly  
 “I confess that I have lied repeatedly.”

Note that in examples (2)–(3) the subordinate clause appears in the subjunctive mood, whereas those in (4)–(5) do not. Both Etxepare (1996) and Broveto (2002) assert that subjunctive mood on the embedded verb is not required for *que*-drop to occur, a claim that is supported with historical and contemporary data from Subirats-Rüggeberg (1987). Broveto argues that other forms of irrealis mood, such as the future or conditional, are also favorable to *que*-drop. On the other hand, Etxepare acknowledges that certain factive verbs in the indicative can allow for *que*-drop, a feature attributable to different underlying syntactic structures for this lexical class of verbs, evidenced by patterns of leftward movement. He observes, along with Torrego (1983), that a pre-verbal embedded subject falling between the matrix clause and the embedded verb will block *que*-drop, possibly even for factive predicates that would otherwise license the Ø option (6) (Etxepare 1998: 473).

- (6) a. *Deseo* Ø *se encuentren mejor*  
 wish.1SG COMP REFL.2P find.2P.SBJV better
- b. \**Deseo* *ustedes* Ø *se encuentren mejor*  
 wish.1SG 2P COMP REFL.2P find.2P.SBJV better  
 “I wish/hope (that) you all feel better.”

Similar observations have been contributed by Antonelli (2012), who analyzes null complementizers in Classical Portuguese, drawing on data from writings dating back

to the 16th and 17th century. He also asserts that the same loosely defined class of lexical verbs – of which each has a cognate in Spanish – are capable of selecting for a null complementizer in their clausal arguments and that this process is blocked by a pre-verbal subject. He adds that a preverbal adverbial may also block *que*-drop.

Also working with historical sources (7)–(9), as well as with print media (10)–(11), Subirats-Rüggeberg (1987: 170–171) presents numerous examples of *que*-drop following verbs associated with volitional/desiderative, deontic and epistemic modality.

- (7) *Eliseo, suplico a Dios Ø abriese los ojos de*  
 Eliseo, beg.1SG to God COMP open.3SG.SBJV the eyes of  
*aquel moço*  
 that boy  
 “Eliseo, I plead to God (that) he open the eyes of that boy.”
- (8) *Quiero, Sancho, Ø me digas lo que acerca de*  
 want.1SG Sancho COMP ACC.1SG tell.2SG.SBJV N.3P COMP related to  
*esto ha llegado a tus oídos*  
 this have.3SG arrived to your ears.  
 “I want, Sancho, (that) you tell me what about this subject has reached your ears.”  
 (Cervantes)
- (9) *Quiera Dios Ø halle algún agradecimiento*  
 want.3SG.SUBJV God COMP find.3SG.SBJV some thanks  
*mi deseo*  
 my wish  
 “God will (that) my wish may find some gratitude.”  
 (Cervantes)
- (10) *Es una profesión que creo Ø puede compararse*  
 COP DET profession COMP think.1SG COMP could.3SG compare-RECP  
*con la suya.*  
 with DET yours  
 “It is a profession which I believe (that) you may compare to yours.”  
 (La Vanguardia 16.12.81: 14.)
- (11) *Pero lo que en principio se pensaba Ø podría*  
 but N.3SG COMP at start PASS think.PST COMP could.COND  
*ser un incidente importante parece que no tendrá*  
 COP DET incident important seem.3SG COMP NEG have.3SG.FUT  
*otras consecuencias*  
 other consequences  
 “But what initially was thought (that) it might be an important incident it seems that it will have no other consequence ...”  
 (La Vanguardia 10.2.81: 49)

Subirats-Rüggeberg goes as far as to argue that this null option is “actually more frequent in colloquial speech than the corresponding sentences where *que* elision has not occurred” (12)–(13) (1987: 168, 169).<sup>2</sup>

- (12) *Esperamos* Ø *sea* *de su* *agrado*  
 hope.1P COMP COP.3SG.SBJV of your pleasure  
 “We hope (that) it please you.”
- (13) *Rogamos* Ø *disculpen* *las molestias*  
 plead.1P COMP excuse.3P.SBJV DET inconveniences  
 “We plead (that) you’ll excuse any inconvenience.”

Thus, we have seen that from the syntactic-theoretical perspective there are several concrete predictions for where null complementizers should appear in Spanish. All the above-referenced authors, save Subirats-Rüggeberg (1987), were careful to emphasize an important caveat: this phenomenon is particular to formal and polite registers. Yet, Silva-Corvalán (1993, 1994, 1998, 2007) attested instances of *que*-drop in elicited narrative interviews of a largely informal nature. She analyzed the speech of first, second and third generation Spanish speakers in Los Angeles. These latter two groups are perhaps less likely to encounter much experience with the formal registers typically associated with *que*-drop. She found that *que*-drop was possible with ‘request clauses’ and ‘estimative verbs’ such as *creer* “to believe”, *pensar* “to think”, and *saber* “to know”, as in (14)–(16), although the only printed examples she gives are with *creer* (Silva-Corvalán 1994, 1998: 137–138). She did not find any examples of *que*-drop in relative clauses or with verbs of saying like *decir* “to say”. Notice that, as opposed to the historical and print examples in (7)–(13), each example in (14)–(16) has a first-person singular referent in the matrix clause accompanied by an overt subject pronoun. This may be indicative of a process of collocation formation or the collapsing of the bi-clausal structure, after which presumably a fixed unit can emerge in specific discursive and syntactic contexts.

- (14) *No muy chiquito, pero yo creo* Ø *ha crecido*  
 NEG very small but SBJ.1SG think.1SG COMP have.3SG grown  
*como – unas seis pulgadas.*  
 like – DET six inches.  
 “Not very small, but I think (that) it has grown about – about six inches.”

---

2. Despite this claim, it appears to me that the examples he provides are rather formal in nature, not colloquial, at least in the modern sense.

- (15) *Yo creo Ø inventaron el nombre que le pusieron.*  
 SBJ.1SG think.1SG COMP invent.3P DET name COMP DAT put.3P  
 “I think (that) they invented the name that they gave her.”
- (16) “*No la hallo muy entusiasmada*”, *yo pensé*  
 NEG ACC.3SG find.1SG very enthusiastic SBJ.1SG think.1.PST  
*entre mí. Yo creo Ø no la quiere ver*  
 between me SBJ.1SG think COMP NEG ACC.3SG want.3SG see  
*[la película] como yo*  
 DET movie like me  
 “‘I don’t think she’s very interested’, I thought to myself. I think (that) she  
 doesn’t want to see [the movie] as much as I.”

While it is tempting to attribute the structure of these utterances to syntactic calquing from English, Silva-Corvalán argues against this conclusion. Instead, she points to attestations of this structure in Eastern Mexican Spanish, where its emergence may have been favored by early contact with indigenous languages that themselves lack obligatory subordinate complementizers (Silva-Corvalán 1998: 230). This conclusion leaves one wanting more, since ultimately it rejects one contact-based explanation for another, equally unsupported one. Seemingly more viable is her argument posed in (1994). There – as part of her broader model for language change during language contact – she explains that since the null complementizer option already existed as a possibility in Spanish, contact with English only serves to promote a “favouring effect in the diffusion of zero *que* in the LA [Los Angeles] bilingual community” (Silva-Corvalán 1994: 138). In this present day comparison of Los Angeles and Mexico City varieties, we could reasonably expect to encounter higher rates of *que*-drop in Los Angeles, be it as a result of direct syntactic calquing or ‘indirect’ contact effects. Before I present the research questions and hypotheses to be tested in the current study, I turn first to a discussion of the research on *that*-drop in the English complementizer system. This should allow for identification of additional syntactic factors that might be relevant for complementizer-drop and also provide an historical example of how a language can undergo a shift from having a complementizer system that seldom allows for null subordinators to one that does so most of the time.

### 3. Subordinating complementizer *that*/Ø in English

Historically, English has undergone a process by which null realizations of the subordinating complementizer *that* have increased dramatically over time. In Old English, overt expression of *that* (*þæt*) was nearly obligatory (Mitchell 1985, cited

in Tagliamonte & Smith 2005). Later during the Middle English period, historical texts like the Wycliffe Sermons reveal overt realizations of *that* at rates as high as 98% (Warner 1982). By the Early Modern period the null option had become much more common, especially with high-frequency verbs such as *know*, *think*, *say* and *tell*, with rates of *that*-drop at around 70% for the period between 1640 and 1710 (Palander-Collin 1997; Rissanen 1991; Warner 1982, cited in Tagliamonte & Smith 2005). Present-day varieties of English demonstrate a sharp contrast between oral and written modes, with *that*-drop occurring at much higher rates in spoken data (Thompson & Mulac 1991a, 1991b; Walker & Cacoullos 2003).

Oral data have revealed strong effects of the lexical matrix verb on the null or overt expression of the complementizer. Thompson & Mulac (1991a, 1991b), in a study conducted at a university in the United States, collected short dialogs between students paired together randomly. They find overwhelming rates of *that*-drop with epistemic verbs, especially collocations such as *I think* and *I guess*, which they label *epistemic phrases*. They attribute this pattern to an ongoing process of grammaticalization by which, because of high token frequency, these collocations underwent semantic bleaching, losing their lexical value in favor of syntactic or discursive-oriented functions. Thus with examples like *I think* and *I guess*, we are not dealing with subordination but rather parentheticals. Thompson & Mulac further demonstrate that higher rates of *that*-drop are also found in specific syntactic contexts: first- or second-person matrix subjects, where no other syntactic elements are to be found in the matrix clause, and when the embedded subject is pronominal.

Thompson & Mulac's account was both corroborated and broadened in a study by Tagliamonte & Smith (2005). The authors compared elicited oral data collected in towns in Northern Ireland, Northern England and Scotland, with findings from previous studies on *that*/Ø from both diachronic and synchronic perspectives. They found similar patterns of *that*/Ø distribution to those found for US American English (Thompson & Mulac 1991a, 1991b) and Canadian English (Walker & Cacoullos 2003), both of which show rates of *that*-drop over 90% in oral modes for frequent matrix verbs like *think*. Other collocation-like subject-verb combinations in the matrix clause were also found to strongly favor *that*-drop. For example, Tagliamonte & Smith (2005) obtained rates of *that*-drop as high as 99% for *you know* and 100% for *I mean*! Besides matrix verb type, other syntactic factors found to be significant in Tagliamonte & Smith include matrix verb tense, the nature of the intervening elements between the matrix and the embedded clause and the position, syntactic category (NP vs. pronoun) and length of embedded subject. Kearns (2007) adds to these factors the person/number features of the embedded subject and co-referentiality between the matrix and embedded subjects.

In order to contextualize their findings in a broader historical perspective, Tagliamonte & Smith (2005) turned to Palander-Collin's (1997) study on the Old English impersonal form *me thinks*. At early stages of development, *me thinks* already permitted *that*-drop even when the option was still scarce in the grammar (Warner 1982). As  $\emptyset$  rates increased over time, *me thinks* began falling into disuse, while *I think* began taking over its function. From there, the syntactic change spread to other parts of the grammar. Other epistemic verbs began to adopt similar conditions for *that*-drop, as eventually did non-epistemic verbs. Thus by connecting the entire trajectory of *that*-drop, Tagliamonte & Smith are successful in showing how a grammaticalization process that began with a disappearing form was able to trigger a far-reaching transformation that impacted the structure of English at numerous levels. Furthermore, they demonstrated that the transformation is not solely the product of grammaticalization targeting the collocation itself. Recall that numerous syntactic factors were found to favor *that*-drop, independently of the lexical status of the matrix verb. This is because the syntactic pattern associated with the collocations – including the status of nearby syntactic elements – is extended to other non-epistemic or non-collocation constructions at varying degrees.

The studies discussed in the preceding paragraphs have set forth some theoretical predictions regarding the distribution of null complementizers in Spanish, some of which have been corroborated by empirical evidence. Additionally, we considered the case of English *that*/ $\emptyset$  distribution as investigated from both diachronic and synchronic perspectives. This provides us with a prime example of how a language can transform from one that almost exclusively requires an overt complementizer to one that permits a null realization the majority of the time. Findings from all of these studies also allow us to make predictions as to what syntactic factors might condition the variable presence of *que* in a variety of Spanish in contact with English, should cross-linguistic syntactic influence be a factor at all. For example, the syntactic-theoretical literature on Spanish would lead us to predict that *que*-drop will occur most often with certain kinds of lexical matrix verbs, including epistemics, desideratives and some factive verbs but not with verbs of saying or in relative clauses. Preverbal subjects in the embedded clause are predicted to block *que*-drop, and it remains unclear whether subjunctive mood on the embedded verb will have any effect at all.

What we know of English *that*-drop confirms that there may be a close relationship between the lexical status of the matrix verbs and the ability to select for a null complementizer, especially with epistemic verbs. We might expect other factors to be important, including any presence of intervening syntactic material between the two clauses, such as adverbials, prepositional phrases, dislocated clauses, objects and clitic pronouns. The status of the subjects of both clauses

may also be at issue, including the person-number combination, co-indexation between the two and their syntactic category, be it pronominal, null or a noun phrase. Each of these factors could potentially prove to influence the distribution of the Spanish null complementizer. If syntactic calquing or borrowing is the source of *que*-drop in Los Angeles Spanish, it should be expected to occur in the same environments as in English. On the other hand, a less direct form of syntactic influence might simply serve to propel the extension or spread of the null complementizer option across different contexts in Spanish, much in the way that the null option began with a very specific context in English and from there spread across the grammar.

So far, the only research that has been carried out on Spanish *que*-drop, both in sociolinguistics and formal linguistics, has looked at a very small number of tokens with complementizer drop. My study builds on this initial sample by carrying out a larger-scale quantitative analysis on un-elicited naturalistically produced data. With respect to linguistic register, since it is known that complementizer drop is already an option in formal registers of Spanish, the present study will focus closely on language produced informally and without elicitation or the presence of the investigator at the time of production. For this reason, I draw my data from social media platform Twitter, a platform used primarily for peer-to-peer communication and the sharing of popular media. Having established a set of preliminary factors that might be relevant for *que*/ $\emptyset$  variation, I turn now to the research questions in the current study and a description of the methodology applied.

#### 4. Goals, research questions, and hypotheses

The current study provides the first quantitative analysis of *que*-drop in two dialects of Spanish: Los Angeles Spanish and Mexico City Spanish. The broad goals of this study are twofold: (i) to assess the status of *que*-drop in Spanish examining both internal and external factors that provide an empirical contribution to the syntactic and contact linguistics literatures; and (ii) to test the efficacy of using social media as a source of data for conducting linguistics research. Bearing in mind these broader goals, the specific research questions that this investigation attempts to answer are as follows:

1. To what extent are null realizations of the subordinating complementizer *que* attested in informal registers of Spanish, such as the informal register employed on social media website Twitter?
  - a. *Que*-drop will be attested in the Twitter corpus though overall rates are predicted to be low.



2. What are the syntactic factors conditioning the distribution of null and overt variants of *que*?
  - a. Predictions put forth in the theoretical syntax literature suggest that *que*-drop should only be possible with certain matrix verbs. This distribution should be contingent upon the modality associated with the matrix verb. Additionally, *que*-drop should be blocked by the presence of an overt post-verbal matrix subject or an overt pre-verbal embedded subject. The syntactic category of the subject is also relevant: noun phrases should disfavor *que*-drop the most, followed by pronominal subjects and then null subjects, which should disfavor *que*-drop the least. Evidence from English *that*-drop suggests that other important factors to consider include the number-person combination of both clausal subjects, where a first-person matrix referent and a second-person embedded referent are most favorable to complementizer-drop. Other factors predicted to block *que*-drop include the presence of intervening material between the two clauses such as objects, prepositional phrases and other adjuncts.
3. Are rates of null complementizers in Los Angeles higher than those found in Mexico City Spanish as a result of contact with English?
  - a. Previous findings from Los Angeles Spanish indicate that higher rates of *que*-drop should be found in a contact variety like Los Angeles Spanish when compared to a mostly monolingual variety like Mexico City Spanish, due to the 'favouring effect' that emerges when parallel structures exist in the languages engaged in a situation of language contact. Direct syntactic calquing should reveal *que*-drop occurring in precisely the same contexts in which *that*-drop occurs while indirect contact effects would result in an expansion of *que*-drop that proceeds in a pattern different from the historical trajectory of English *that*-drop.

As highlighted above, another indirect goal of this study is to assess the utility of extracting data from online social media for the purpose of linguistic structural analysis. The pragmatic and discursive properties of language use on Twitter are not yet fully understood, but generally the linguistic register used on the media platform is mostly informal in nature since character count limitations encourage brevity and tweets are generally used for personal communication and the sharing of popular culture media.

I turn now to a description of the methodology adopted including the procedure for data extraction. Since this method is somewhat new, at least for the purposes of linguistic-structural analysis, I will make brief mention of other studies that have also applied this procedure.



## 5. Methodology

Using online social media as a source for quantitative data has a strong tradition in several social sciences, most notably the field of business economics (Pang & Lee 2008). In the field of linguistics there has been more hesitation to embrace Twitter as a source for language data, though in the past five years several studies have proven successful. Much of the initial work has focused on the discursive properties of Twitter language. For example, Zappavigna (2011) built a corpus using a data-scraping script in Python that collected evaluative tweets expressing opinions on President Barack Obama in the 24 hours following the 2008 US elections. Her focus was on how power relationships and political affiliation are negotiated in this still new medium of communication.

Other studies have begun to use Twitter as simply another source of data for the analysis of the structural properties of language, rather than exploring the discursive features of the Twitter platform itself. Haddican & Johnson (2012) consider regional effects on Particle Verb Alternation in Southern England, Scotland and North America. To address potential dialectal differences in word order, they compared a controlled judgment task with a Twitter corpus compiled with data extracted over the course of 4 months. Their use of *geolocation* – the ability to specify a region by longitude, latitude, and radius – allowed them to conduct a successful comparison of syntactic dialectal variation in each region. Christodulelis (2014) also used data drawn from Twitter to explore the syntactic factors conditioning the use of inflected infinitives in Brazilian Portuguese. Gonçalves & Sánchez (2014) collected a large corpus of Tweets from varieties of Spanish across the globe. They sorted lexical items into various conceptual categories associated with dialectal variation in Spanish such as “popcorn”, “car”, and “computer”. By applying a machine-learning algorithm, they were able to identify two distinct ‘Superdialects’ of Spanish. These Superdialects appear to be more closely related to population density rather than nation-state boundaries. For example, Superdialect  $\alpha$  was representative of large urban centers in the Americas and Spain while Superdialect  $\beta$  was representative of rural, more conservative dialects of Spanish on both sides of the Atlantic.

Clearly the potential benefits of Twitter-drawn data are numerous for linguistics, especially for the application of corpus-based methodologies. Needless to say, the sheer size of the body of data it produces is impressive, with an average of 500 million tweets being produced worldwide every day (Kirkorian 2013)! It has an advantage over data collected via survey in that tweets are un-elicited, produced naturally and willingly, and as such remove some of the biases associated with interview or survey-based methodologies (McCormick et al. 2013). Additionally, studies in natural language processing have developed

metrics to determine the degree of formality associated with a particular social media outlet or forum. Collecting data from a number of different social media forums and websites in Spanish, Mosqueda & Moreda (2012) were able to show that Twitter, along with other platforms that place limits on character count like Kongregate, are associated with a more informal speech style and feature the frequent use of slang, colloquialisms, innovative uses of spelling and non-orthographic symbols and creative grammatical structures. One drawback to this methodology is the difficulty of accessing sociolinguistic and language background, as well as biographical information. Information posted on a user's profile may be inaccurate or absent all together and is currently not easily extracted. Additionally, when using geolocation, there is no way to be certain that all the data extracted from a given location will contain tweets produced only from users who actually reside in that region. Somebody visiting Mexico City for the weekend may be sending out tweets, and this may appear in the researcher's dataset unwittingly. Nevertheless, the relative ease of access and exorbitant amounts of data makes for a tempting tool to use. So long as its limitations are acknowledged, it should prove useful for the large-scale analysis of structural linguistic features.

With this in mind, we turn now to the procedures implemented in the current study. All data were extracted using the software application Plus One Social Insight (POSI henceforth; Husting 2014). As a private entity, Twitter is able to regulate access by paid services to messages, or *tweets*, sent using their website. The Twitter API is the interface between Twitter's infrastructure and the World Wide Web through which applications like POSI can extract tweets. POSI will have an interface familiar to users of any search engine. One can query a particular user's Twitter *handle*, or username, a topic or string known as *hashtag* marked by the symbol # or simply a string of text including regular typed messages, along with hyperlinks or symbols such as *emoticons*. To illuminate, a search for the hashtag *#complementizers* will return any messages marked as such by a user within the text of tweet. Similarly, a search for a particular string of typed graphemes such as *creo* will return all tweets containing this string. All searches on POSI return hits only from tweets made in the prior 6 to 7 days. All tweets are then linked to a Microsoft Excel 2013 spreadsheet via a bridge file in Microsoft Access 2013 where PowerPivot charts provide extensive meta-data, mostly concerned with the density of the network of messengers, trending topics and the device being used to send the tweet. Only those users whose privacy and sharing settings allow for geolocation will return on a given query. For the current study, queries were returned from two different regions, a 50 mile radius around Mexico City at 34.040029 N, 118.246 W (Figure 1) and a 50 mile radius around Los Angeles, California at 19.430075 N, 99.137 W (Figure 2).



Figure 1. Area of extraction, Mexico City



Figure 2. Area of extraction, Los Angeles

Data were extracted on three separate occasions over the course of a year. The first extraction was performed on December 28, 2013. Five lone matrix verbs, all first-person singular in the present indicative, were queried: *espero*, *creo*, *juro*, *digo*, *confieso*. Together the queries returned ~65,000 tweets, of which 800 were included in the final sample. The second extraction was performed on August 22, 2014, producing ~15,000 tweets. This time, the set of verbs queried was expanded greatly and multiple queries were performed to capture each

number-person instantiation in the present indicative and also the preterit and imperfect past tense forms (Appendix A). Of this set, approximately 170 tweets were retained for the final analysis. The final extraction was performed on January 3, 2015, producing ~35,000 tweets. This was done in the same manner as the previous extraction, with numerous matrix verbs in various person-number-tense instantiations. 535 additional tweets were pulled from this sample. Ultimately 1,505 tweets were retained for analysis. Sections of each dataset extracted were chosen randomly by using the randomized order function in the matrix verb column, an operation for randomly sorting a column in Excel 2013. Each section was scanned for the manual removal of all retweets and direct quotations. *Retweets* are tweets selected by other users and subsequently forwarded as messages on their own Twitter handle. In the database they are conspicuously marked with the letter combination RT and are thus easy to identify and eliminate if necessary. Since direct quotations also do not reflect unique, individual language production, these too were deleted wherever evident. Any syntactic structures outside the envelope of variation were also excluded. For example, any instance of matrix verb + non-finite verb (i.e. *quiero estar*) was eliminated, among others.

For binary dependent factor ‘complementizer status,’ levels were ‘null’ or ‘overt’. In order to be coded as ‘overt,’ a complementizer could be realized with the following graphemes or combination of graphemes: *que*, *ke*, *qe*, *k*, *q*, and even *y*.<sup>3</sup> Among the independent factors, six were linguistic and two were extra-linguistic. Each linguistic factor concerned the syntactic and morphological elements in proximity to the verbs in each clause. The type and position of the subject in the matrix clause was categorized as ‘null,’ ‘pre-verbal pronoun,’ ‘post-verbal pronoun,’ ‘pre-verbal noun phrase,’ ‘long pre-verbal noun phrase’ (four words or more), ‘post-verbal noun phrase,’ or ‘long post-verbal noun phrase.’ Another factor, the ‘subject of the embedded clause,’ was coded identically. For the multivariate analysis, these categories were collapsed into three categories – ‘null,’ ‘pre-verbal’ and ‘post-verbal’ – due to low levels of variation across cells.

In Thompson & Mulac (1991a) and Kearns (2007), it was shown that the person-number combination of the two clausal subjects was important in influencing rates of null complementizers. For this reason, a factor for each possible person-number combination was considered. This factor eventually proved not to demonstrate sufficient variability across cells and was excluded from the

---

3. Observe the use of *y* in the following tweets: @SaraPrz *suenaba bastante bien... necesito el permiso del workkk pero ya ando en eso, espero y me lo den!!*@ekbuk *Que bien, de vacaciones...?? Espero y las disfrutes mucho...!! Cuidate...!! las disfrutes mucho...!! Cuidate...!!*

multivariate statistical analysis. However, the various levels were collapsed into one binary factor that indicated whether there was ‘co-indexation’ between both clausal subjects or ‘no co-indexation’ when there was none.

There were conflicting claims in the theoretical literature as to the relationship between irrealis mood and *que*-drop (cf. Brovetto 2002 vs. Etxepare 1998 vs. Subirats-Rüggeberg 1987). For this reason, the ‘grammatical mood of the embedded verb’ was taken into consideration and could be coded as ‘indicative’, ‘conditional’ or ‘subjunctive’.

To determine the effects of other ‘syntactic elements that may intervene between the matrix verb and embedded verbs’, the following elements were considered: ‘negation’, ‘adverbials’, ‘clitic pronouns’, ‘other’, ‘multiple elements’, or ‘none’. The ‘other’ category included prepositional phrases, dislocated clauses and any other type of adjunct or non-clausal elements.

For the factor ‘matrix verb’, each verb was classified according to the kind of modality it expressed (Appendix 1). These categories were ‘volitional’, ‘epistemic’, ‘verbs of saying’, and ‘stative verbs’. These latter two categories were collapsed into a single category.

Two extra-linguistic factors were included in the analysis. The continuous factor ‘word count’ was included to ensure that *que*-drop was not simply a reflex of Twitter’s 140 character count limit. So as to address the question of potential cross-linguistic influence in the complementizer system of Los Angeles Spanish, ‘city’ was included as a factor with levels ‘Mexico City’ and ‘Los Angeles’. In order to highlight the coding process, (16) and (17) contain examples of tweets and their corresponding classification. Table 1 summarizes each factor with their respective levels.

Table 1. Factors and levels

Co-reference	City	Number/person subjects combination	Character count	Complementizer status (DV)
Yes	Los Angeles	(eliminated)	(continuous)	Overt
No	Mexico D.F.		1–140	Null
Embedded subject	Matrix subject	Intervening material	Mood	Matrix verb modality
Pre-verbal	Pre-verbal	Negation	Subjunctive	Epistemic
Post-verbal	Post-verbal	Adverbial	Conditional	Volitional
Null	Null	Clitic	Indicative	Stative-Saying
		None		
		Multiple		
		Clause/other		

- (16) *Espero que los #ReyesMagos me traigan*  
 hope.1SG COMP DET #ThreeKings ACC.1SG bring.3P.SBJV  
*algo. :) something ☺*  
 “I hope that the Three Kings bring me something ☺”

Coding scheme:

Embedded subject – *pre-verbal NP*  
 Matrix subject – *null*  
 Intervening material – *clitic*  
 Mood – *subjunctive*  
 Matrix verb modality – *volitional*  
 Co-reference – *none*  
 City – *Mexico City*  
 Character count – 46

- (17) *Espero mi mamá haya preparado algo sabroso*  
 hope.1SG my mom have.3SG.SBJV prepared something tasty  
*hoy para comer! ☺👉 today to eat! ☺👉*  
 “I hope my mom has prepared something tasty to eat today! ☺👉”

Coding scheme:

Embedded subject – *pre-verbal NP*  
 Matrix subject – *null*  
 Intervening material – *none*  
 Mood – *subjunctive*  
 Matrix verb modality – *volitional*  
 Co-reference – *none*  
 City – *Mexico City*  
 Character count – 63  
 Complementizer – *null*

## 6. Results

In this section, I begin with descriptive observations and then proceed to the inferential analysis. The distribution of matrix verb modality in the final sample is shown in Figure 3. Epistemic verbs (*creer* “to think”, *pensar* “to think”, *suponer* “to suppose”, *dudar* “to doubt”, *imaginar* “to imagine”, *saber* “to know”) were most frequent in the corpus at 46%, most of these from the verb *creer*. Volitional verbs (*esperar* “to hope”, *querer* “to want”, *desear* “to wish”, *rogar* “to plead”, *pedir* “to ask for”) were the second most frequent type at 32% and were dominated by *esperar*.

Verbs of saying and stative verbs (*decir* “to say”, *ver* “to see”, *confesar* “to confess”, *lamentar* “to lament”, *prometer* “to promise”) represented 22% of all tokens in the corpus with *decir* being the most common verb of this category.

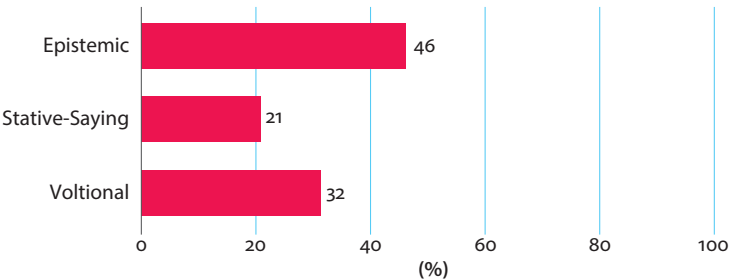


Figure 3. Portion of corpus by matrix verb modality

I now move to the distribution of null complementizers in relation to each of the syntactic factor groups. Table 2 shows the total number of tokens for each modality type with overt and null complementizers, as well as the percentage of null complementizers in the total sample represented by each modality. The rightmost column shows the percentage of null complementizers within each modality category.

Table 2. Distribution in the sample of null complementizers by modality type

Modality	Overt	Null	Proportion of all null	Proportion within category null
Volitional	354	134	79%	27%
Stative-Saying	306	16	9%	5%
Epistemic	675	19	12%	3%
Total	1335	169	13% (of total sample)	-

Figure 4 provides a visualization of the distribution of the null complementizer according to the modality of the matrix verb. Below it, Figure 5 shows the percentage within modality category realized with a null complementizer.

The nature of the embedded subject was also predicted to have an effect on *que*/Ø distribution. For Spanish, Etxepare (1996) suggests that overt pre-verbal subjects in the complement clause block the realization of a null complementizer as a consequence of interrupted adjacency. Meanwhile in English, pronominal embedded subjects favored *that*-drop while noun phrases, and especially noun phrases that are three words or longer, disfavored the Ø option. Thus we would predict the following hierarchy with the favorability of *que*-drop decreasing from left to right (Figure 6).



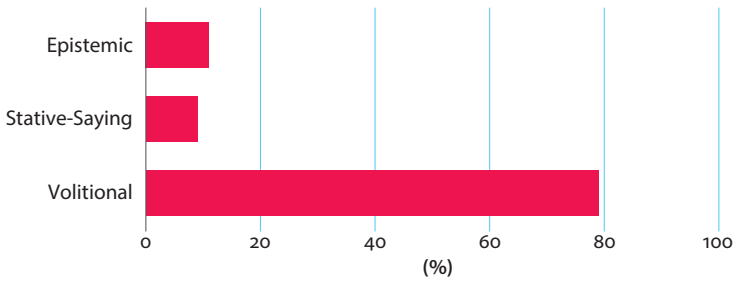


Figure 4. Portion of total sample Ø COMP by modality type

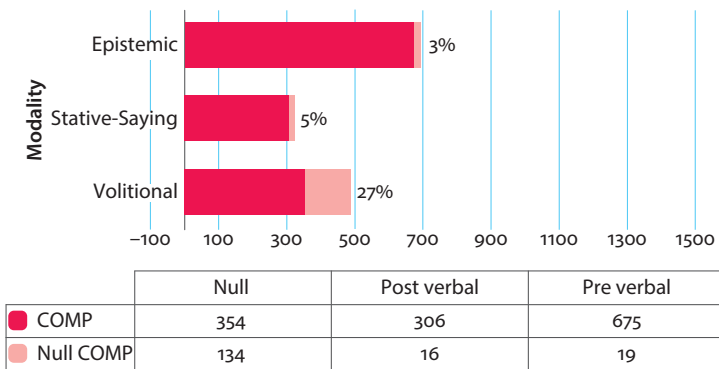


Figure 5. Portion of Ø COMP within each modality category

Null > Post-verbal pronoun = Post-verbal NP > Pre-verbal pronoun > Pre-verbal NP  
 more *que*-drop less *que*-drop

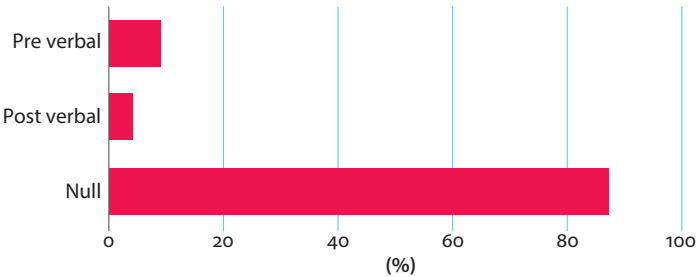
Figure 6. Predicted hierarchy of *que*-drop by embedded subject type/position

This hierarchy is partially confirmed in Table 3. Tokens with null embedded subjects contributed 87% of all null tokens in the sample as shown in Figure 7, and of all tokens containing a null embedded subject, 15% also contained a null complementizer. The subject type/position that contributed the second most null complementizers were pre-verbal subjects at 9% followed by post-verbal subjects at 4%. Within categories, 6% of all tokens with a pre-verbal subject also lacked *que* while 9% of all tokens with a post-verbal subject did as well. A visualization of the within-category proportions is shown in Figure 8. In Table 4 I show the distribution of *que*-drop for modality type and embedded subject type/position together. In this way, we can see how the combination of a volitional verb combined with an embedded null subject is most highly favorable to *que*-drop.

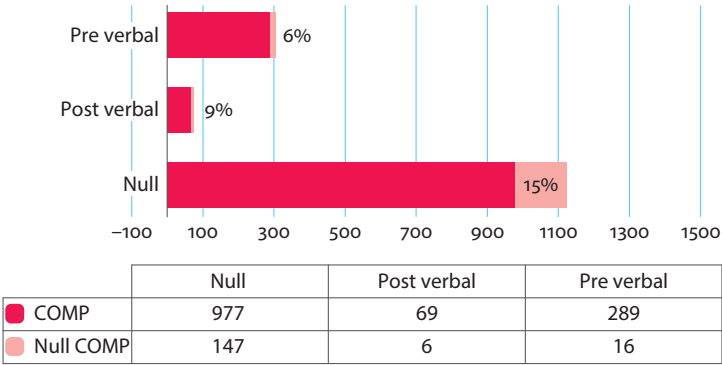


**Table 3.** Distribution of null complementizers in the sample by embedded subject type/position

	Overt COMP	Null COMP	Proportion of all null COMP	Proportion within category null COMP
<i>Null subject</i>	977	147	87%	15%
<i>Post-verbal</i>	69	6	4%	9%
<i>Pre-verbal</i>	289	16	9%	6%
<i>Totals</i>	1335	169	11% (of total)	



**Figure 7.** Portion of total sample Ø COMP by embedded subject type/position



**Figure 8.** Portion of tokens Ø COMP within each embedded subject type/position category

Factors that did not ultimately emerge as statistically significant included the ‘type and position of the matrix subject’, ‘co-reference between the two subjects’, ‘grammatical mood of the embedded verb’, ‘intervening syntactic material’, ‘character count of tweets lacking a complementizer’, and ‘city’.

To further explore the influence of factors on *que*-drop, I conducted a binomial logistic regression using a step-up/step-down procedure with the package Rbrul (Johnson 2008) for the programming language R (R Core Team 2015). Application

value was ‘null’ for the dependent factor ‘complementizer status’. All independent variables are included in Table 1. The factor ‘character count’ was included as a continuous factor, and Rbrul was prompted to test for interactions between factors ‘modality type: irrealis’. The results of the analysis are shown in Table 5.

**Table 4.** The distribution of null complementizers according to both modality and the status of the embedded subject

	Null subj. + null COMP	Post-verbal + null COMP	Pre-verbal + null COMP
<i>Volitional</i>	117	4	10
<i>Stative-Saying</i>	13	1	2
<i>Epistemic</i>	15	1	3

**Table 5.** Binomial logistic regression application value ‘null’ for factor ‘complementizer status’

Modality matrix verb ( <i>p</i> < .005)	Log odds	Tokens	Response proportion	Factor weight
<i>Volitional</i>	1.522	488	0.275	0.821
<i>Stative-Saying</i>	−0.466	322	0.050	0.386
<i>Epistemic</i>	−1.056	694	0.027	0.258
				Range = 0.563
<b>Embedded SUBJ (<i>p</i> &lt; .005)</b>				
<i>Null</i>	0.442	1124	0.131	0.609
<i>Post</i>	0.072	75	0.080	0.518
<i>Pre</i>	−0.513	305	0.052	0.374
				Range = 0.235
Deviance	861.252			
Mean	0.112			
Df	5			
R <sup>2</sup> Nagelkerke	0.313			

I had postulated that the modality of the matrix verb would be the strongest predictor of *que*-drop, with higher rates predicted to be associated with volitional and epistemic modality. Stative verbs and verbs of saying were predicted to disfavor *que*-drop. Modality was confirmed as the strongest predictor of *que*-drop overall (*p* < .005). Volitional verbs indeed strongly favored *que*-drop (FW = 0.821); however, epistemic verbs disfavored the null option (FW = 0.258). The same was true for verbs of saying/stative verbs, which disfavored *que*-drop slightly less (FW = 0.386).

Theoretical accounts of null complementizers in Spanish predict that an overt pre-verbal subject in the embedded clause would block the  $\emptyset$  option (Antonelli 2012; Brovett 2002; Etxepare 1998; Torrego 1983). This prediction was mostly borne out in the data, with the type and position of the embedded subject also achieving significance ( $p < .005$ ). Null embedded subjects (FW = .609) as well as overt post-verbal embedded subjects (FW = .518) favored the null option while pre-verbal subjects disfavored (FW = .374) this option.

The factor ‘city’, which had been included to test for differences in the expression of the null complementizer in Los Angeles and Mexico City Spanish, did not achieve significance. To further explore if any differences arise in the hierarchy of factors conditioning the distribution of *que*-drop across these two dialect regions, each set of data was separated for individual statistical analysis with 706 tokens from Mexico City and 793 tokens from Los Angeles. Two separate binomial logistic regressions were conducted using Rbrul (Johnson 2008), and they reveal an identical distribution of factors in both samples. The results are shown for Mexico City in Table 6 and for Los Angeles in Table 7.

**Table 6.** Binomial logistic regression application value ‘null’ for factor ‘complementizer’, Mexico City sample

Modality matrix verb ( $p < .00005$ )	Log odds	Tokens	Response proportion	Factor weight
<i>Volitional</i>	1.798	230	.317	0.858
<i>Stative-Saying</i>	−0.816	148	0.034	0.307
<i>Epistemic</i>	−0.983	327	0.028	0.272
				Range = .586
<b>Embedded SUBJ (<math>p &lt; .005</math>)</b>				
<i>Null</i>	0.562	529	0.146	0.637
<i>Post</i>	0.059	32	0.094	0.515
<i>Pre</i>	−0.621	144	0.049	0.349
				Range = 0.288
Deviance	403.724			
Mean	0.123			
Df	5			
R <sup>2</sup> Nagelkerke	0.371			

These results reveal a strikingly similar distribution in both grammars with respect to *que*-drop. Table 8 shows the raw numbers and percentages associated with each city with respect to modality type side-by-side for convenience.

**Table 7.** Binomial logistic regression application value ‘null’ for factor ‘complementizer’, Los Angeles sample

Modality matrix verb ( <i>p</i> < .005)	Log odds	Tokens	Response proportion	Factor weight
<i>Volitional</i>	1.315	257	.233	0.788
<i>Stative-Saying</i>	−0.269	169	0.059	0.433
<i>Epistemic</i>	−1.046	366	0.027	0.26
				Range = 0.528
<b>Embedded SUBJ (<i>p</i> &lt; .05)</b>				
<i>Null</i>	0.441	591	0.118	0.608
<i>Post</i>	0.133	43	0.070	0.533
<i>Pre</i>	−0.574	158	0.044	0.36
				Range = 0.248
Deviance	439.68			
Mean	0.101			
Df	5			
R <sup>2</sup> Nagelkerke	0.28			

**Table 8.** Side-by-side comparison of null complementizers by modality type in Mexico City and Los Angeles

	Los Angeles				Mexico City			
	Overt COMP	Null COMP	% within % null	% within category null	Overt COMP	Null COMP	% within % null	% within category null
<i>Volitional</i>	197	60	75%	23%	157	73	84%	32%
<i>Stative-Saying</i>	159	10	13%	6%	143	5	6%	3%
<i>Epistemic</i>	356	10	12%	3%	318	9	10%	3%

This side-by-side comparison shows only a slight expansion of *que*-drop in Los Angeles Spanish, where verbs of saying/stative verbs contributed 13% and epistemic verbs contributed 12% to the total count of null complementizers compared to 6% and 10%, respectively, in Mexico City. Next, §7 includes a discussion of these results.

7. Discussion

In this section I address each of the research questions posed in §4, while also framing the results within the broader context of language variation and change.

Our first question asked to what extent null complementizers are a feature of these varieties of Spanish, at least within the informal and non-normative registers found on Twitter. The overall rate of null complementizers with respect to the entire size of the sample analyzed is relatively low: only 169 tokens out of 1,505, or 11%, were realized without *que*. Nonetheless, this remains a notable amount and probably more than previously would have been expected from an informal register of Spanish or from intuitive guesses based on oral language modes.

The question of Twitter's character count limitation remains: could speakers be dropping *que* as a strategy for keeping messages short and concise? This appears not to be the case. Remember that there are other shortening strategies available to the user: many speakers realize *que* as *qe* or *ke*, or as a single letter *k*, *q* or *y*. This would allow them to keep their complementizers overt while only taking up one character's worth of space. Of course any other words in the clause could also be targeted for deletion or shortening. Still, to be fully confident that this was not the case, character count was included as a continuous factor in the multivariate analysis. It did not emerge as a significant predictor of *que*-drop ( $p > .05$ ). In addition, the average character count for all tweets missing *que* was only 93.7, well below the limit of 140. Thus we conclude that while still a somewhat infrequent occurrence, *que*-drop is indeed an option following certain predicates, even in informal registers in both Los Angeles Spanish and Mexico City Spanish.

My second research question was concerned with the syntactic elements in the clause that might influence *que*-drop. Drawing on observations from theoretical syntax, combined with evidence from one sociolinguistic study and several studies on English *that*, we were able to formulate clear-cut predictions as to what structural factors might condition the distribution of *que*-drop. Our strongest prediction was for the modality of the matrix verb. Claims from formal syntax varied as to their classification of each verb with respect to its modality or lexical value (cf. Brovetto 2002 vs. Etxepare 1998), but it would seem that null complementizers are related to modality in some way and could appear following a matrix clause containing an epistemic, desiderative, volitional or deontic verb. This is partially confirmed in Silva-Corvalán's data. She finds that *que*-drop occurred with the epistemics *creer*, *pensar* and *saber*. On the other hand, verbs of saying and stative verbs like *decir* or *ver* were predicted to disallow *que*-drop. Similar observations have been made for Classical Portuguese (Antonelli 2012), and studies on English *that*-drop have also demonstrated ample evidence that the lexical class of the matrix verb exerts great influence on the ability to drop the complementizer in the following clause, at least to the extent that this feature interacts with other elements in the clause. Epistemic verbs like *think*, *know*, *mean* and *guess* were found to trigger *that*-drop the most (Tagliamonte & Smith 2005).

The present findings show that volitional verbs, especially tokens with the matrix verb *esperar*, were by far the most common context in which *que*-drop occurred. Other verb types contributing the next highest rates of *que*-drop were epistemic verbs like *creer*, which contributed 9% of all null realizations of *que*, along with verbs of saying/stative verbs, with *jurar* and *decir* each contributing 5%. The results for verbs of saying/stative verbs are probably the most surprising result here since we would have expected verbs of saying and stative verbs to disallow *que*-drop altogether, while epistemics would be expected to be more permissive of *que*-drop. As a modality class, verbs of saying/stative verbs disfavored *que*-drop slightly less than epistemics, even though *creer* was shown to select for a null option more frequently than other non-volitional verbs. So clearly, in both English and Spanish, there is a strong link between modality and the status of the complementizer. In English, null realizations of *that* are most strongly connected to epistemic modality, but this factor was shown to work in tandem with several syntactic factors to determine the ultimate *that*/ $\emptyset$  distribution (Kearns 2007; Tagliamonte & Smith 2005; inter alia). In Spanish, the null option is more closely associated with volitional modality, while lower rates of *que*-drop can still be found with epistemic verbs and with verbs of saying/stative verbs.

Following the modality of the matrix verb, the next most relevant factor was the status and position of the embedded subject. This category sought not only to determine how the position of the subject with respect to the verb in the complement clause affected the distribution of null *que* but also to test for the syntactic category of the subject, or in other words, whether the subject was a noun phrase, a pronoun or unrealized. A number of generative studies (Antonelli 2012; Brovetti 2002; Etxepare 1998; Torrego 1983) have suggested that the subject has to be post-verbal or null, so that nothing would intervene between the two clauses, in order for the null complementizer to proceed. In English, since subjects can only be pre-verbal and are usually required to be phonetically realized, the subjects with the least phonetic content, pronouns, are those that most strongly promote *that*-drop (Tagliamonte & Smith 2005). The predictions are clear then: post-verbal subjects, be they noun phrases or pronouns, along with null realizations of the embedded subject, should be most favorable to *que*-drop. Least favorable should be a pre-verbal NP, with a pre-verbal pronoun being only slightly more permissive of *que*-drop (Figure 6). This hierarchy plays out almost exactly, with both null and post-verbal embedded subjects favoring *que*-drop. Thus we see that as with English, the modality of the matrix verb and the status of the embedded verb combine to exert an influence on the distribution of *que*-drop.

One of the more surprising results was that the factor 'city' failed to achieve significance. Mexico City actually had a higher proportion of *que*-drop at 12.5%, compared to Los Angeles' 10%. This, therefore, provides evidence in support of

Silva-Corvalán's (1993, 1994, 1998, 2007) observation that *que*-drop has long been an available option in monolingual Spanish, and its presence in Los Angeles Spanish should not be classified as a recent syntactic borrowing. In fact, empirical results seem to indicate that if English is exerting any significant syntactic influence, it is merely to broaden the range of contexts in which the null option may occur, since as can be inferred from Table 8 *que*-drop in Los Angeles Spanish appears to be employed more frequently with stative verbs/verbs of saying and epistemic modality types than in Mexico City Spanish. Note, however, that this difference did not achieve statistical significance.

So rather than a process of syntactic calquing, we instead are seeing a shift in the complementizer system of both Los Angeles and Mexico City Spanish. It is evident that the pragmatic domain of *well-wishing* and the *expression hope and desire* – domains usually identified as pertaining to the realm *volitional modality*, or sometimes more specifically identified as *boulomaic modality* (Kiefer 1998; Rescher 1968) – interacts with the syntactic context and frequency effects to target *routinization* on a particular verb, *esperar*. Recall that two of the three extractions of tweets were around the New Year and Christmas holidays. Since *esperar* was the verb form with the highest rates of null complementizers, I looked more closely at several of the tweets containing this matrix verb type. Many of the tweets lacking *que* came from tokens that used *esperar* to express *well-wishing* or to articulate one's hopes and desires for the benefit and well-being of the recipient, as in (18)–(21).

- (18) @JuanPachuco<sup>4</sup> *Hola hermosos, espero Ø estés de*  
 @JuanPachuco hello beautiful.P hope.1SG COMP COP.2SG.SBJV in  
*lujo vamos al gotcha!*  
 luxury go.IMP.1P to-DET gotcha  
 “@JuanPachuco Hello handsome ones, I hope you're living itup! Let's go to Gotcha!”  
 (Mexico City)
- (19) @PekeñaCM *sanguijuelita esperó Ø estés muy bien un*  
 @PekeñaCM little leech hope.1SG COMP COP.2S very well DET  
*abrazote a ti y a la princesa*  
 big hug to you and to DET princess  
 “@PekeñaCM little-leech I hope you're doing very well a big hug to you and the princess”.  
 (Los Angeles)

---

4. To maintain the anonymity of the Twitter users, I have anonymized all Twitter handles.

- (20) @AvSanchez *Espero* Ø *termines pronto y tengas*  
 @AvSanchez hope.1SG COMP finish.2SG soon and have.2SG.SBJV  
*una agradable noche campeón. Un placer saludarte*  
 DET enjoyable night champion DET pleasure greet-ACC  
 “@AvSanchez I hope you finish soon and have an enjoyable evening, champ.  
 It’s a pleasure to greet you!” (Mexico City)
- (21) *Espero con ansias el 2015 🙏 espero* Ø  
 await.1SG with anxiousness DET 2015 🙏 hope.1SG COMP  
*sea un año lleno de bendiciones. De mi parte sólo*  
 COP.3SG.SBJV DET year full of blessings. From my part only  
*habrá trabajo. ☺*  
 be.3SG.FUT work ☺  
 “I anxiously await 2015. I hope it will be a year full of blessings. As for me,  
 there will only be work ☺.” (Los Angeles)

The other context that dominated the use of *esperar* involves the expression of a desire for a particular real-world outcome on the part of the person producing the tweet, on the part of his or her interlocutor or on the part of some third party, as in (22)–(25).

- (22) @Rodricito *esperamos* Ø *nos mandes foto de*  
 @Rodricito hope.1P COMP ACC.1P send.2SG.SBJV photo of  
*tu visita!*  
 your trip  
 “@Rodricito we hope you send us photos from your trip!”
- (23) *Si de mi lado te vas espero* Ø *no*  
 COND from my side RELF go.2SG hope.1SG COMP NEG  
*regreses mas.*  
 return.2SG.SBJV more  
 “If you leave my side I hope you never come back again.”
- (24) *Buen partido de Chivas espero* Ø *sea un buen*  
 good game of Chivas hope.1SG COMP COP.3SG.SBJV DET good  
*torneo #DaleRebaño*  
 tournament #mopthefloorwiththem  
 “Good game by Chivas I hope it will be a good tournament #DaleRebaño”
- (25) #ExBeliebers *Espero* Ø *este reto no afecté la*  
 #ExBeliebers hope.1SG COMP this challenge NEG affect.3SG.SBJV DET  
*amistad entre Beliebers y Directioners*  
 friendship between Beliebers and Directioners  
 “#ExBeliebers I hope this challenge doesn’t affect the friendship between  
 Beliebers and Directioners”



So, it appears that this particular pragmatic context – that of *well-wishing* or the expression of the *desire for a beneficial real-world outcome* – is strongly associated with the null complementizer option. While some tweets involved overlap between these two modes, the majority could be classified confidently as one or the other. A close analysis of all tokens with *que*-drop and with matrix verb *esperar* revealed that 44% involved well-wishing, while 56% involved the expression of a hope or desire for a particular real-world outcome to occur, not related to the holiday well-wishing.

This kind of formulaic construction is often associated with *routinization*, where a process of *grammaticalization* can result in the reduction of morpheme boundaries and a collapsing of bi-clausal structures (Hopper & Traugott 1993: 72–73). Indeed the process by which the  $\emptyset$  option for English *that* became so prevalent is widely agreed to have emerged from a process of *grammaticalization* as a result of *routinization*. Thompson & Mulac (1991a, 1991b) argue that high rates of *that*-drop can be attributed to grammaticalization of matrix clauses with *I think* and *I guess* undergoing a reduction in bi-clausal structure so that they have transformed into parentheticals, rather than verbs that select for a clausal complement. Tagliamonte & Smith (2005) argue that grammaticalization worked in tandem with other principles of language change to shape the historical trajectory of *that*/ $\emptyset$  distribution. They point to the Middle English impersonal interjection *me thinks*, a form that allowed for high rates of null complementizers even at a stage when English accepted almost exclusively overt complementizers (Warner 1983). This form eventually merged with *I think*, which quickly underwent a process of grammaticalization where it solidified into a collocation of sorts, admitting rates of null complementizers in the 90% range for spoken language (Tagliamonte & Smith 2005, and citations therein). This process also triggered a massive shift in the grammar, allowing for *that*-drop with other verb forms, first in the same syntactic contexts as with collocations such as *I think*, *I guess* or *you know*, but later spreading to other contexts as well. Even in synchronic data, the interplay between the lexical value or modality of the matrix verb and other grammatical elements is evident (Tagliamonte & Smith 2005; Walker & Cacoullos 2003).

To determine if the modality of the matrix verbs only triggers *que*-drop in specific syntactic constructions, a closer analysis of just the tokens with matrix *esperar* and lacking *que* proved necessary. If in fact structures with matrix *esperar* were beginning to form a collocation that indexes a particular modality, we should see that this occurs only in a specific syntactic context where the person-number realization of the matrix and embedded subjects are consistent across tokens and the clausal complexity is minimal in the clauses adjacent the subordinator in question (Traugott 2005). For tokens with matrix *esperar*, the pattern is undeniable: in 99% of clauses with matrix *esperar* and also lacking *que*, the matrix referent is in

first person. In 39% of these, the embedded referent is in third person and in 57% of these, the embedded referent is in second person. It becomes clear that routinization is facilitated by syntactically predictable contexts: the matrix referent is virtually always in first person, while the embedded referent is in second person most often but also third person a notable amount of the time. Similarly, matrix *esperar* is never realized in a tense or mood other than the present indicative when followed immediately by an embedded clause. Observe, then, the creation a true collocation of the structure *espero* + *second- or third-person referent*. This collocation only emerges in contexts of volitional and boulomaic modality and indexes well-wishing or the desire for a beneficial real-world outcome in a given situation. With respect to clausal complexity, we find not a single token with matrix *esperar* and *que*-drop where there is an overt matrix subject, and only one token with matrix negation and one token with a matrix adverbial. Thus, the process of routinization and the collapse of the bi-clausal structure is free to proceed, absent the clausal complexity that usually hinders the formation of a collocation as a result of grammaticalization (Traugott 2005).

In Spanish we see a process that is less far-reaching and less dramatic than what we have seen historically for English *that*. *Que*-drop has not become tied to a larger process of grammaticalization that would trigger a massive shift in the entire complementizer system. Instead, the syntactic domain in which *que*-drop is permitted has likely remained relatively stable over time and is only now extending gradually into this highly specific context with matrix *espero*. If pure frequency were the dominant influencing factor in this grammaticalization process, one would expect more frequent verbs like *creer* and *decir* to demonstrate rates of *que*-drop along the lines of *esperar*. While clearly they do admit this option, the rate at which *que*-drop occurs with *esperar* is dramatically higher.

It should not be surprising that complementizer expression would communicate distinctions in modality. The left periphery has long been known to be the locus of the syntax-pragmatics interface (Bahtchevanova 2007; Rizzi 1997; *inter alia*), and several languages have complementizers that index the modality associated with a matrix clause (Bahtchevanova 2007; Frajzyngier 1995), as has been previously claimed for Spanish *que* (Murphy 2008; Pons Bordería 2003; Porroche Ballesteros 2000). I hypothesize that the null option in subordinating structures is associated with volitional and boulomaic modality, especially well-wishing and the communication of hope and desire for a beneficial outcome. *Esperar* is frequently used in this specific pragmatic context and often so when in the syntactic context of number-person combination *Matrix 1P + Embedded 2P/3P*. This creates a syntactic environment that is predictable. What is more, clausal complexity is minimal since other non-core elements are omitted from the matrix clause. As a result of the confluence of these syntactic and pragmatic-discursive elements,

*espero*P has become the target of routinization, thus resulting in the beginnings of the formation of collocation, calling into question whether these are subordinating structures at all or, similarly to *I think* or *you know* in English, operate more like parentheticals.

There may also be historical parallels with the trajectory of the complementizer system in English. Much as English *that*-drop had its roots in a historical form *me thinks*, it is possible that the process associated with *que*-drop following *esperar* has a similar historical origin – albeit a less impactful one – in a form that shares the semantic domain of hope and desire and thus also indexes volitional and boulomaic modality. The interjection *ojalá* ‘I hope’ via Mozarabic from the Arabic الله شاء, *law sha’a Allah*; *si Dios quisiera* ‘If God would will it so’ / ‘If god so wills it’, clearly shares a semantic affinity with *esperar*, and this form has long been known to optionally appear with complementizer *que* followed by a clause in the subjunctive mood (Zagona 2002: 53). Furthermore, it was shown in Subirats-Rüggeberg (1987) that historically, and in press media, the zero option has long been viable with numerous other matrix verb types from several other semantic domains. So, we know the structure has lingered in the grammar of Spanish for a long time. Additionally, we know that *que*-drop was permissible in Classical Portuguese with the same verb types as Spanish (Antonelli 2012), and this language shares the interjection *oxalá* of the same origin as the Spanish lexeme.

Now, for Spanish, we may be witnessing an expansion of the context in which *que*-drop can occur. This expansion is strongly tied to the aforementioned specific pragmatic context and thus shows some evidence of routinization with *esperar*. For verbs other than *esperar*, such as verbs of saying, stative verbs and other deontic, epistemic or desiderative verbs, *que*-drop continues to be an option, although it occurs at low overall rates. This distribution has probably remained stable over the years, as evidenced by the presence of this structure in the historical literature (Subirats-Rüggeberg 1987) and in another empirical study (Silva-Corvalán 1994). In this way, it may reflect a prototypical case of ‘stable variation’ as Fafulas et al. (this volume) found for *pa/para* variation in Venezuelan Spanish; however, more sociolinguistic and historical evidence than was provided here may be necessary to confirm this assertion for *que*-drop. The expansion of the zero option for *esperar* found in the current study may signal a more recent shift in the distribution of the null complementizer, where this option is largely becoming associated with this particular matrix verb in the right syntactic-pragmatic context. More diachronic analyses coupled with analysis of *que*-drop in other dialectal regions will be necessary to reveal the details of the evolution of such a process over time and space.

Here, I have shown that *que*-drop is a viable option in the grammars of Mexico City and Los Angeles Spanish that is conditioned by syntactic and lexical factors such as the modality of the matrix verb and the status and position of the embedded

subject. Finding few differences between the distributions of factors conditioning *que*-drop in each region, I reject the hypothesis that Los Angeles Spanish would employ this structure because of direct borrowing from English syntax. Rather, I hypothesize that English is either having limited impact upon Los Angeles Spanish in this regard or serving to expand the contexts in which the already viable zero option is possible (cf. Silva-Corvalán 1994). Los Angeles exhibited slightly higher rates of  $\emptyset$  than Mexico City with epistemics and verbs of saying/stative verbs, and this may be a clue that a generalization of the favorable context for  $\emptyset$  is underway. This would be consistent with similar findings for Spanish in contact with English. For example, Orozco (2015) finds expanded use of the analytic future at the expense of the periphrastic future in New York City Puerto Rican and Colombian Spanish in contexts where the latter is generally preferred in non-contact varieties. He demonstrates that the historical trajectory of the periphrastic vs. analytic future has been characterized by much variation in Spanish and that the current patterns of usage have been driven primarily by grammaticalization. When in contact with English, this historical process is extended so that the preferred form in the superstrate language, the analytic future, becomes generalized even further in the substrate language. Thus rather than direct syntactic borrowing, only indirect contact effects can be observed: the analytic future emerges as the preferred form in NYC Spanish as a result of 'indirect' syntactic influence rather than direct syntactic calquing or borrowing. Kyzar (this volume) arrived at strikingly similar conclusions in his analysis of future tense expression in the Spanish of residents of Mexican descent residing in Baton Rouge, Louisiana. He too finds an increase in the proportion of periphrastic future used to express futurity, at the expense of the simple present and especially the morphological (analytic) future, in the youngest members of the speech community. His conclusions support those of Orozco, pointing to an extension of an ongoing historical change in the grammar of Spanish accelerated by language contact and predicting the eventual obsolescence of the morphological future.

Recall however that in both dialects we find that *que*-drop is strongly favored by constructions with the matrix verb *esperar* and particularly in contexts of expressing wishes for the well-being of the interlocutor. In lieu of sufficient evidence supporting a contact-driven change, we are forced to turn to explanations grounded in language-internal phenomena, in this case the routinization of a syntactic chain. This routinization is facilitated by a syntactically simple and predictable configuration and might be on the way to the collapse of the subordinating structure into a true collocation. Due to this fact, and due to what we know about the diachronic evolution of *law sha'a Allah* into *ojalá*, I speculate that the null complementizer in Spanish appears to index the expression of volitional and boulomaic modality. Further historical analysis should assess if high rates of *que*-drop with *esperar* has

resulted from semantic affinity with *ojalá*, which performs an identical pragmatic function and has also been subjected to syntactic routinization.

Last but not least, this investigation also makes a methodological contribution to the study of language variation. The use of corpora constructed from Twitter has the potential to dramatically change the state of the art in corpus linguistics. Not only can they provide us with copious amounts of data that are easily accessible, free of the observer's paradox, and mostly informal in register; they have also proven to be particularly useful for the analysis of forms that are relatively rare in the input and difficult to elicit in an interview or laboratory setting.

## References

- Antonelli, André L. 2013. "The Structure of Complementizerless Clauses in Classical Portuguese". *Romance Linguistics 2012: Selected Papers from the 42nd Linguistic Symposium on Romance Languages (LSRL), Cedar City, Utah, 20–22 April 2012*, ed. by Jason Smith & Tabea Ihsane, 185–198. Amsterdam & Philadelphia: John Benjamins.
- Bahtchevanova, Mariana. 2007. "Mood, Modality, and Complementation: A Cross-Linguistic Study of the Syntax and Semantics of the Left Periphery". Tempe, Ariz.: Ph.D. dissertation, Arizona State University.
- Brovetto, Claudia. 2002. "Spanish Clauses without Complementizers". *Current issues in Romance languages* ed. by Teresa Satterfield, Christina Tortora & Diana Cresti, *Current Issues in Linguistic Theory* vol. 220, 19–43. Amsterdam & Philadelphia: John Benjamins.  
doi:10.1075/cilt.220.04bro
- Cattell, Ray 1978. "On the Source of Interrogative Adverbs". *Language* 54.61–77.  
doi:10.2307/412999
- Christodoulelis, Eleni. 2014. *The Importance of Clause Type in Portuguese (Un)Inflected Infinitives. Paper presented at New Ways of Analyzing Variation (NWAY) 43, October 11, 2014. Chicago.*
- Elseness, Johan 1984. "That or Zero? A Look at Choice of Object Clause Connective in a Corpus of American English". *English Studies: A Journal of English Language* 65.519–33.  
doi:10.1080/00138388408598357
- ¿Es peligroso fumar en un Taxi? (1977, Feb. 1), ABC. Retrieved from abc.es.
- Etxepare, Ricardo. 1996. "On Ø Complementizers in Spanish". *Anuario del Seminario de Filología Vasca Julio de Urquijo* 30:2.469–496.
- Etxepare, Ricardo. 1998. "On the Structure of Declarative Clauses". *Romance Linguistics: Theoretical Perspectives* [Current Issues in Linguistics Theory vol. 160] ed. by Armin Schwegler, Bernard Tranel & Myriam Uribe Etxebarria, 53–66. Amsterdam: John Benjamins.
- Frajzyngier, Zygmunt. 1995. "Functional Theory of Complementizers". *Modality in Grammar and Discourse* ed. by Joan Bybee & Suzanne Fleischman, 473–502. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/tsl.32.21fra
- Gonçalves, Bruno & David Sánchez. 2014. "Crowdsourcing Dialect Characterization through Twitter". *PLOS One*. Retrieved from <http://journals.plos.org/plosone/article?id=10.1371/journal.pone011207>. doi:10.1371/journal.pone.0112074

- Haddican, Bill & Daniel Ezra Johnson. 2012. "Effects on the Particle Verb alternation across English Dialects". *University of Pennsylvania Working Papers in Linguistics* 18.2. Retrieved from <http://repository.upenn.edu/pwpl/vol18/iss2/5>.
- Hopper, Paul J. & Elizabeth C. Traugott. 1993. *Grammaticalization*. Cambridge: Cambridge University Press.
- Husting, Patrick. 2014. *Plus One Social Insight*. <http://plusonesocial.com/>.
- Johnson, Daniel Ezra. 2008. "Getting off the GoldVarb Standard: Introducing Rbrul for Mixed-Effects Variable Rule Analysis". *Language and Linguistics Compass* 3:1.359–383. doi:10.1111/j.1749-818X.2008.00108.x
- Kearns, Kate. 2007. "Epistemic Verbs and Zero Complementizer". *English Language and Linguistics* 11.475–505. doi:10.1017/S1360674307002353
- Kiefer, Ferenc. 1998. "Modality". *Handbook of Pragmatics* 4 ed. by Jef Verschueren, Jan-Ola Östman, Jan Blommaert & Chris Bulcaen, 1–36. Amsterdam: Benjamins.
- Kirkorian, Raffi. 2013, August 16. *New Tweets per Second Record, and How!* Retrieved from <https://blog.twitter.com/2013/new-tweets-per-second-record-and-how>.
- McCormick, Tyler, Hedwig Lee, Nina Cesare & Ali Shojaie. 2013. "Using Twitter for Demographic and Social Science Research Tools for Data Collection". *CSSS Working Papers*, 127. Seattle, Wash.: University of Washington Center for Statistics and the Social Sciences.
- Mitchell, Bruce. 1985. *Old English Syntax*, vols. 1–2. Oxford: Clarendon Press. doi:10.1093/acprof:oso/9780198119357.001.0001
- Mosqueda, Alejandro & Paloma Moreda. 2012. "SMILE: An Informality Classification Tool for Helping to Assess Quality and Credibility in Web 2.0 Texts". *AAAI Technical Report WS-12-02 Real-Time Analysis and Mining of Social Streams*.
- Murphy, Melissa. 2008. *The Role of Typological Drift in the Evolution of the Romance Subjunctive: A Study in Word-Order Change, Grammaticalization, and Synthesis*. Austin, Tex.: Ph.D. dissertation, University of Texas at Austin.
- Orozco, Rafael. 2015. "Castilian in New York City: What Can We Learn From the Future?" *New Perspectives on Hispanic Contact Linguistics in the Americas* ed. by Sandro Sessarego & Melvin González-Rivera, 347–371. Madrid & Frankfurt: Iberoamericana/ Vervuert.
- Palander-Collin, Minna. 1997. "A Medieval Case of Grammaticalization, *methinks*". In *Grammaticalization at Work* (Studies of Long-Term Developments in English 24) ed. by Matti Rissanen, Merja Kytö & Kirsi Heikkonen, 371–403. Berlin & New York: Mouton de Gruyter. doi:10.1515/9783110810745.371
- Pang, Bo & Lillian Lee. 2008. "Opinion Mining and Sentiment Analysis". *Foundations and Trends in Information Retrieval* 2:1–2.1–135. doi:10.1561/15000000011
- Pons, Bordería, Salvador. 2003. "*Que* inicial átono como marca de modalidad". *Estudios de lingüística* 17.531–545.
- Porroche Ballesteros, Margarita. 2000. "Algunos aspectos de *que* en el español conversacional". *Círculo de lingüística aplicada a la comunicación* 3. <http://www.ucm.es/info/circulo/no3/porroche.htm>.
- R Development Core Team 2008. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org>.
- Rescher, Nicholas. 1968. *Topics in Philosophical Logic*. Dordrecht: Reidel. doi:10.1007/978-94-017-3546-9
- Rissanen, Matti. 1991. "On the History of that/zero as Object Clause Links in English". *English Corpus Linguistics: Studies in Honour of Jan Svartvik* ed. by Karin Aijmer & Bengt Altenberg, 272–89. London & New York: Longman.



- Rizzi, Luigi. 1997. "The Finite Structure of the Left Periphery". *Elements of Grammar* ed. by Lillian Haegeman, 281–337. Dordrecht: Kluwe. doi:10.1007/978-94-011-5420-8\_7
- Silva-Corvalán, Carmen. 1993. "On the Permeability of Grammars: Evidence from Spanish and English Contact". *Linguistic Perspectives on Romance Languages* ed. by William J. Ashby, Marianne Mithun & Giorgio Perissinotto, 19–43. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/cilt.103.08sil
- Silva-Corvalán, Carmen. 1994. *Language Contact and Change: Spanish in Los Angeles*. Oxford: Oxford University Press.
- Silva-Corvalán, Carmen. 1998. "On Borrowing as a Mechanism of Syntactic Change". *Romance Linguistics: Theoretical Perspectives* ed. by Armin Schwegler, Bernard Tranel & Myriam Uribe-Etxebarria, 225–247. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/cilt.160.18sil
- Silva-Corvalán, Carmen. 2007. "The Limits of Convergence on Language Contact". *Language Contact and the Dynamics of Language: Theory and Implications*. Symposium conducted at the Max Planck Institute Leipzig.
- Subirats-Rüggeberg, Carlos. 1987. "Sentential Complementation in Spanish: A Lexico-Grammatical Study of Three Classes of Verbs". *Linguisticae Investigationes Supplementa: Études en Linguistique Française et Générale* 14 ed. by Jean Claude Chevalier, Maurice Gross & Christian Leclère. Amsterdam & Philadelphia: John Benjamins.
- Tagliamonte, Sali & Jennifer Smith. 2005. "No Momentary Fancy! The Zero 'Complementizer' in English Dialects". *English Language and Linguistics* 9.289–309. doi:10.1017/S1360674305001644
- Thompson, Sandra A. & Anthony J. Mulac. 1991a. "A Quantitative Perspective on the Grammaticalization of Epistemic Parentheticals in English". *Grammaticalization II* ed. by Elizabeth C. Traugott & Bernd Heine, 313–339. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/tsl.19.2.16tho
- Thompson, Sandra A. & Anthony J. Mulac. 1991b. "The Discourse Conditions for the Use of the Complementizer *that* in Conversational English". *Journal of Pragmatics* 15.237–251. doi:10.1016/0378-2166(91)90012-M
- Traugott, Elizabeth C. 2005. "Constructions in Grammaticalization". *The Handbook of Historical Linguistics* ed. by Brian Joseph & Richard Janda, 624–647. Oxford: Wiley-Blackwell.
- Torrego, Esther. 1983. "More Effects of Successive Cyclic Movement". *Linguistic Inquiry* 14.561–565.
- Villa-García, Julio. 2014. *The Syntax of Multiple-Que Sentences in Spanish: Along the Left Periphery*. Amsterdam & Philadelphia: John Benjamins. doi:10.1075/ihl.2
- Walker, James & Rena Torres Cacoullos. 2003. *How to Take a Complement in Canadian English*. Paper presented at New Ways of Analyzing Variation (NWAV) 32, October 2003. Philadelphia.
- Warner, Anthony. 1983. *Complementation in Middle English and the Methodology of Historical Syntax*. University Park, PA: Penn State University Press.
- Zagona, Karen. 2002. *The Syntax of Spanish*. Cambridge: Cambridge University Press.
- Zappavigna, Michele. 2011. "Ambient Affiliation: A Linguistic Perspective on Twitter". *New Media Society* 13.788. doi:10.1177/1461444810385097

## Appendix: All verb types in sample

### Stative/Verbs of saying

*Ver*

*Jurar*

*Decir*

*Confesar*

*Prometer*

*Lamentar*

### Epistemic

*Suponer*

*Saber*

*Pensar*

*Imaginar*

*Dudar*

*Creer*

### Volitional

*Rogar*

*Querer*

*Esperar*

*Desear*

*Pedir*





# Index

## A

Academia Norteamericana de la Lengua Española 28  
 Afro-Costa Rican women 38, 40–42  
 Afro-Ecuadorian Spanish 63–66  
 Afro-Hispanic 64–66, 68, 70  
 Afro-Peruvian Spanish 74, 95  
 Amazonian Spanish 180–182, 190  
 Andean Spanish 22, 193–194  
 argumentation theory 158–59, 165–167  
 assimilation 14–16, 36  
 attrition 64, 71, 73, 75, 76

## B

bilingualism 16, 36, 39, 49, 71, 89, 112, 203, 204, 257  
 borrowing 111–113  
 boulomaic modality 301, 324, 327–329  
 British Council 19, 20  
 business communication 16, 128, 134, 135, 145

## C

case distinction 74, 92, 99, 175–180, 189–193  
 causality 153–154, 157–158  
 cyclicity 271  
 change in progress 225–227  
 Caraqueño Spanish 225, 229, 230, 232  
 Chinchano Spanish 85, 95, 103  
 Classical Portuguese 302, 322, 328  
 collocation 300, 304, 306, 326–329  
 Colombian Spanish 176–177, 180  
 competition 92, 179, 192, 194, 225, 252–253  
 Creole 42, 43, 70–72, 86, 101  
 commissive 127–147  
 cyberspace 11, 12, 26, 27, 30

## D

discourse marker 111–113, 153–156

## E

epistemic modality 303, 319, 323, 324  
 expletive *que* 200–202

## F

face 18, 129, 131, 139, 146  
 female migrations 36, 37, 40, 41  
 future tense 137, 138, 251–257, 302, 329

## G

gender 42, 131, 207–210, 224, 238, 259, 294  
*see also* grammatical 40, 68, 69, 73–75, 104, 175, 180, 185, 189–191  
 globalization 11, 12, 14–16, 19, 20, 25, 26, 29, 30  
 grammaticalization 265, 270, 301, 306, 326

## I

interface hypothesis 71–74, 76  
 interference 200, 203, 206,  
 Istanbul 64–66

## J

Judeo-Spanish 63–77

## L

L2 Spanish 63, 75, 103, 132, 180, 181, 281, 283  
 language  
*see also* as commodity 12, 16  
 attitudes 203–206, 228, 257, 283  
 maintenance 35, 42, 57  
 minoritized 11, 12, 19, 22, 23, 27, 30  
 normalization 11, 12, 23, 25, 29

power in 17, 19, 36, 128, 135, 140, 141, 310  
 prestige in 200, 204, 240  
 shift 36, 42, 57, 65–67, 225–227, 255, 270, 305, 324  
 late capitalism 11–13, 16, 17, 30  
*leísmo* 178–181, 193  
 Limón Province 37–41  
 locative prepositions 200, 201  
 Los Angeles 29, 300, 305, 308, 311, 314  
 Louisiana Spanish 128, 134–135, 147, 194, 249–251, 258–260

## M

marketplace economy 11, 12, 16, 19, 27, 30  
 mass media 11, 12, 25, 28–30  
 matched-guise technique 199, 200, 204, 205  
 Mexico City 13, 29, 249, 305, 311, 314, 323  
*Miami Herald* 14, 15  
 Minimalism 86, 89, 92, 98, 104, 105  
 motivation 36, 104, 161, 179, 279, 280  
*see also* heritage learner 287  
 instrumental 280–282  
 integrative 280–282  
 intrinsic 280, 282

## N

'neutral' Spanish 29, 31, 204  
 New Orleans 135  
 null subject 72, 73, 85–90, 92, 103, 317–321  
 Null Subject Parameter 72, 73, 85–87, 103  
 number (grammatical) 67, 69, 73–75, 99, 104, 177, 180, 185, 260, 308, 313

## O

Observer's Paradox 67, 226, 240, 330

overt subject 63, 67, 69, 70, 72,  
73, 76, 85, 87–96, 103, 105,  
251, 316, 320

## P

*pa(ra)* alternation 224–228  
partitive *de* 200–203  
phi-agreement 63, 67–69, 71,  
76, 92  
politeness 19, 130–132, 135, 142  
postmodernity 11, 12, 14, 22,  
26, 28, 30  
Principle of distance 253  
*pro*-drop 73, 89, 90, 92, 95, 103  
pronoun 18, 63, 70, 73, 76, 87–91,  
175–76, 201–202, 313–14, 323  
*Proyecto 3399* 22, 24

## R

routinization 324, 326–330

## S

Sephardic 64, 65  
simplification 181, 194, 257  
socioeconomic status 38, 204,  
207, 224–226, 230, 259, 260  
solidarity 16, 19, 59, 64, 131,  
200, 205, 206  
speech act 127–140, 157  
stable variation 224, 225, 227

## T

temporal distance 255, 266  
*that*-drop 300, 301, 305–307,  
322, 328

Tikuna 176–177, 181–182

Tojol-ab'al 111–113

*tuteo* 11, 12, 16–18, 21

Twitter 300, 309–311, 313,  
322, 330

## U

United Fruit Company 37  
Universal Grammar 75, 86,  
89, 98

## V

Valencian 200–203  
vernacular 13, 36, 64, 215, 226,  
227  
volitional modality 301,  
323, 324

This collection of original contributions dealing with Hispanic contact linguistics covers an array of Spanish dialects distributed across North, South, and Central America, the Caribbean, the Iberian Peninsula, and the Bosphorus. It deals with both native and non-native varieties of the language, and includes both synchronic and diachronic studies. The volume addresses, and challenges, current theoretical assumptions on the nature of language variation and contact-induced change through empirically-based linguistic research. The sustained contact between Spanish and other languages in different parts of the world has given rise to a wide number of changes in the language, which are driven by a concomitance of different linguistic and social processes. This collection of articles provides new insight into such phenomena across the Spanish-speaking world.

"This volume, which includes empirically-based studies conducted on Spanish varieties across the Americas, the Caribbean, the Iberian Peninsula and the Bosphorus, constitutes an important contribution to the literature on Hispanic contact linguistics, enriching our understanding of variation and change in the Spanish language in the postmodern era."

**Carol Klee**, *University of Minnesota*

"This collection provides important new pieces to the puzzle of variation around the Spanish-speaking world, including macro and micro analyses as well as monolingual and bilingual contexts. Broad in scope and high in quality."

**Kim Potowski**, *University of Illinois at Chicago*

"The contributions to this volume represent the state of the art in the study of variation in Spanish and the effects of contact on Spanish dialects, and King and Sessarego have done a masterful job in compiling such a strong collection. It will be a point of reference for scholars and graduate students for many years to come."

**Scott Schwenter**, *The Ohio State University*

ISBN 978 90 272 0014 3



**JOHN BENJAMINS PUBLISHING COMPANY**